

# Peaks and dips in Gaussian random fields: a new algorithm for the shear eigenvalues, and the excursion set theory

Graziano Rossi<sup>1,2,3\*</sup>

<sup>1</sup> *Korea Institute for Advanced Study, Hoegiro 85, Dongdaemun-Gu, Seoul 130 – 722, Korea*

<sup>2</sup> *CEA, Centre de Saclay, Irfu/SPP, F-91191 Gif-sur-Yvette, France*

<sup>3</sup> *Paris Center for Cosmological Physics (PCCP) and Laboratoire APC, Université Paris 7, 75205 Paris, France*

Accepted 2012 November 27. Received 2012 November 7; in original form 2012 June 29

## ABSTRACT

We present a new algorithm to sample the constrained eigenvalues of the initial shear field associated with Gaussian statistics, called the ‘peak/dip excursion-set-based’ algorithm, at positions which correspond to peaks or dips of the correlated density field. The computational procedure is based on a new formula which extends Doroshkevich’s unconditional distribution for the eigenvalues of the linear tidal field, to account for the fact that halos and voids may correspond to maxima or minima of the density field. The ability to differentiate between random positions and special points in space around which halos or voids may form (i.e. peaks/dips), encoded in the new formula and reflected in the algorithm, naturally leads to a straightforward implementation of an excursion set model for peaks and dips in Gaussian random fields – one of the key advantages of this sampling procedure. In addition, it offers novel insights into the statistical description of the cosmic web. As a first physical application, we show how the standard distributions of shear ellipticity and prolateness in triaxial models of structure formation are modified by the constraint. In particular, we provide a new expression for the conditional distribution of shape parameters given the density peak constraint, which generalizes some previous literature work. The formula has important implications for the modeling of non-spherical dark matter halo shapes, in relation to their initial shape distribution. We also test and confirm our theoretical predictions for the individual distributions of eigenvalues subjected to the extremum constraint, along with other directly related conditional probabilities. Finally, we indicate how the proposed sampling procedure naturally integrates into the standard excursion set model, potentially solving some of its well-known problems, and into the ellipsoidal collapse framework. Several other ongoing applications and extensions, towards the development of algorithms for the morphology and topology of the cosmic web, are discussed at the end.

**Key words:** methods: analytical – methods: statistical – galaxies: formation – galaxies: halos – cosmology: theory – large-scale structure of Universe.

## 1 INTRODUCTION

The large-scale spatial organization of matter in clusters, filaments, sheets and voids, known as the *cosmic web* (Peebles 1980; Bardeen et al. 1986; Bond et al. 1991), is the manifestation of the anisotropic nature of gravitational collapse. This typical filamentary pattern has been confirmed by observations, for example with the 2dFGRS, SDSS and 2MASS redshift surveys of the nearby Universe (Colless et al. 2003; Tegmark et al. 2004; Huchra et al. 2005), and is routinely seen in large scale  $N$ -body numerical simulations – see for example the New Horizon Runs (Kim et al. 2011) or the recent Millennium-XXL (Angulo et al. 2012). Basic characteristics of the cosmic web are the spatial arrangement of galaxies and mass into elongated filaments, sheet-like walls and dense compact clusters, the existence of large near-empty void regions, and the hierarchical nature of this

\* E-mail: graziano.rossi@cea.fr

mass distribution (Aragón-Calvo et al. 2007, 2010a,b; Zhang et al. 2009). In particular, as pointed out by Bond, Kofman & Pogosyan (1996), ‘embryonic’ cosmic web is already present in the primordial density field. These key properties, along with the alignment of shape and angular momentum of objects (i.e. Argyres et al. 1986; Catelan et al. 2001; Lee & Springel 2010), are mainly due to the effects of the tidal field, associated with the gravitational potential – while the Hessian of the density field (i.e. the inertia tensor) plays a secondary role in determining the characteristic pattern of the cosmic web.

Pioneering works devoted to the key role of the initial tidal field in shaping large-scale structures trace back to Doroshkevich & Zeldovich (1964), Doroshkevich (1970), Zeldovich (1970), Sunyaev & Zeldovich (1972), Icke (1973), and Doroshkevich & Shandarin (1978); their studies have contributed to reach a solid understanding of the formation and evolution of structures. Other classical works in this direction (i.e. Peebles 1980; White 1984; Bardeen et al. 1986; Kaiser 1986; Bertschinger 1987; Dubinski 1992; Bond & Myers 1996; Bond, Kofman & Pogosyan 1996; van de Weygaert & Bertschinger 1996) have considerably improved the statistical description of the cosmic web from first principles. Their impact is broader, as there is a correspondence between structures in the evolved density field and local properties of the linear tidal shear; this allows one to estimate the morphology of the cosmic web (Bond & Myers 1996; Rossi, Sheth & Tormen 2011), and is crucial in understanding the nonlinear evolution of cosmic structures (Springel et al. 2005; Shandarin et al. 2006; Pogosyan et al. 2009), the statistical properties of voids (Lee & Park 2006; Platen, van de Weygaert & Jones 2008; Lavaux & Wandelt 2010), the alignment of shape and angular momentum of halos (West 1989; Catelan et al. 2001; Faltenbacher et al. 2009; Lee & Springel 2010), and for characterizing the geometry and topology of the cosmic web (Gott et al. 1986, 1989; Park & Gott 1991; Park et al. 1992, 2005; van de Weygaert & Bond 2008; Forero-Romero et al. 2009; Aragon-Calvo et al. 2010a,b; Choi et al. 2010; Shandarin et al. 2010; van de Weygaert et al. 2011; Cautun et al. 2012; Hidding et al. 2012). In addition, the eigenvalues of the mass tidal tensor can be used to classify the large-scale environment (Shen et al. 2006; Hahn et al. 2007; Zhang et al. 2009) – hence their fundamental importance.

In the context of the initial shear field, Doroshkevich (1970) provided a key contribution by deriving the joint probability distribution of an ordered set of eigenvalues in the tidal field matrix corresponding to a Gaussian potential, given the variance of the density field. Throughout the paper, we will refer to it as the *unconditional* probability distribution of shear eigenvalues. Because the initial shear field associated with Gaussian statistics plays a major role in the formation of large scale structures, considerable analytic work has been based on Doroshkevich’s unconditional formulas since their appearance, but those relations neglect the fact that halos (voids) may correspond to maxima (minima) of the underlying density field. The local extrema of such field (i.e. peaks/dips) are plausible sites for the formation of nonlinear structures (Bardeen et al. 1986), and some numerical studies have indeed reported a good correspondence between peaks in the initial conditions and halos at late times (see in particular Ludlow & Porciani 2011). Hence, their statistical properties can be used to predict the abundances and clustering of objects of various types, and in studies of triaxial formation of large-scale structures (Bardeen et al. 1986; Bond & Myers 1996).

Motivated by these reasons, recently Rossi (2012) has provided a set of analytic expressions which extend the work of Doroshkevich (1970) and Bardeen et al. (1986) – and are akin in philosophy to that of van de Weygaert & Bertschinger (1996). These new relations incorporate the peak/dip constraint into the statistical description of the initial shear field, and are able to differentiate between random positions and peak/dips in the correlated density field. In essence, they allow one to express the joint probability distribution of an ordered set of eigenvalues in the initial shear field, given the fact that positions are peaks or dips of the density – and not just random spatial locations. These relations are obtained by requiring the density Hessian (i.e. the matrix of the second derivatives of the density field, associated with the curvature) to be positive/negative definite, which is the case in the vicinity of minima/maxima of the density. The correlation strength between the gravitational and density fields is quantified via a reduced parameter  $\gamma$ , which plays a major role in peaks theory (i.e.  $\gamma$  is the same as in Bardeen et al. 1986 for Gaussian smoothing filters – see Appendix A). Doroshkevich’s (1970) *unconditional* formulas are then naturally recovered in the absence of correlation, when  $\gamma = 0$ . From these new *conditional* joint probabilities, it is possible to derive the individual distributions of eigenvalues subjected to the extremum constraint, along with some other related conditional probabilities; their expressions were also provided in Rossi (2012), extending the work of Lee & Shandarin (1998).

The primary goal of this paper is show how the main conditional formulas presented in Rossi (2012) and reported here (Equations 1 and 11) lead to a new algorithm – called the ‘peak/dip excursion-set-based’ algorithm – to sample the constrained eigenvalues of the initial shear field associated with Gaussian statistics, at positions which correspond to peaks or dips of the correlated density field. While it is clearly possible to sample the constrained eigenvalues of the tidal field directly from the conditional probability distribution function, as done for example by Lavaux & Wandelt (2010) in the context of cosmic voids, the theoretical work carried out by Rossi (2012) allows for a much simpler procedure, part of which was previously thought not to be achievable analytically (see again the Appendix B in Lavaux & Wandelt 2010). Besides providing novel theoretical insights, the main strength of the algorithm resides in its natural inclusion into the standard excursion set framework, allowing for a generalization. As we will discuss in Section 6, this technique potentially solves a long-standing problem of the standard excursion set theory; moreover, it is well-suited for the ellipsoidal collapse framework.

The other goal of the paper is to present a first physical application of the new algorithm, related to the morphology of the cosmic web. In particular, we provide a new expression for the conditional distribution of shape parameters (i.e. ellipticity and prolateness) in the presence of the density peak constraint, which generalizes some previous literature work and combines the formalism of Bardeen et al. (1986) – based on the density field – with that of Bond & Myers (1996) – based on the shear field. The formula has important implications for the modeling of non-spherical dark matter halos and their evolved halo shapes, in relation to the initial shape distribution. In addition, we also test and confirm our theoretical predictions for the individual distributions of eigenvalues subjected to the extremum constraint, along with other directly related conditional

probabilities. Finally, we illustrate how this algorithm can be readily merged into the excursion set framework (Peacock & Heavens 1990; Bond et al. 1991), and in particular to obtain an excursion set model for peaks and dips in Gaussian random fields. The key point is the ability to differentiate between random positions and peaks/dips, which is contained in Equations (1) and (11) and encapsulated in the algorithm. While we leave this latter part to a dedicated forthcoming publication, we anticipate the main ideas here. We also discuss several other ongoing applications and extensions, towards the development of algorithms for classifying cosmic web structures.

The layout of the paper is organized as follows. Section 2 provides a short review of the key equations derived in Rossi (2012), which constitute the theoretical framework for the new algorithm described here; the main notation adopted is summarized in Appendix A, for convenience. Moving from this mathematical background, Section 3 presents the new ‘peak/dip excursion-set-based’ algorithm (some insights derived from this part are left in Appendix B), while Section 4 tests its performance against various analytic distributions of eigenvalues and related conditional probabilities, subjected to the extremum constraint. Section 5 shows a physical application of the computational procedure, towards the morphology of the cosmic web. A new expression for the conditional distribution of ellipticity and prolateness in the presence of the density peak constraint is also given; in particular, the description of Bardeen et al. (1986) is combined with that of Bond & Myers (1996). Section 6 illustrates how this new algorithm can be readily used to implement an excursion set model for peaks and dips in Gaussian random fields, and makes the connection with some previous literature. Finally, Section 7 discusses several ongoing and future promising applications, which will be presented in forthcoming publications, and in particular the use of this algorithm for triaxial models of collapse and in relation to the morphology and topology of the cosmic web.

## 2 JOINT CONDITIONAL DISTRIBUTION OF EIGENVALUES IN THE PEAK/DIP PICTURE

We begin by reexamining two main results derived in Rossi (2012), which constitute the key for developing a new algorithm to sample the constrained eigenvalues of the initial shear field. This part may be also regarded as a compact review of the main formulas for the constrained shear eigenvalues, which can be used directly for several applications related to the cosmic web. The notation adopted here is the same as the one introduced by Rossi (2012), with a few minor changes to make the connection with previous literature more explicit. It is summarized in Appendix A; a reader not familiar with the notation may want to start from the appendix first. In particular, in what follows we do not adopt ‘reduced’ variables, so that the various dependencies on  $\sigma$  values (i.e.  $\sigma_T \equiv \sigma_0$  for the shear, and  $\sigma_H \equiv \sigma_2$  for the density Hessian) are now shown explicitly. However, for the sake of clarity, we omit to indicate the understood dependence of these two global parameters in the left-hand side of all the formulas. Note that we prefer to write  $\sigma_T$  and  $\sigma_H$ , rather than  $\sigma_0$  and  $\sigma_2$ , for their more intuitive meaning (i.e. the label  $T$  indicates that a quantity is connected to the shear field, while the label  $H$  points to the Hessian of the density field). As explained in Appendix A, we also introduce the 6-dimensional vectors  $\mathbf{T} = (T_{11}, T_{22}, T_{33}, T_{12}, T_{13}, T_{23})$  and  $\mathbf{H} = (H_{11}, H_{22}, H_{33}, H_{12}, H_{13}, H_{23})$ , derived from the components of their corresponding shear and density Hessian tensors.

Rossi (2012) obtained the following expression for the probability of observing a tidal field  $\mathbf{T}$  for the gravitational potential, given a curvature  $\mathbf{H}$  for the density field:

$$p(\mathbf{T}|\mathbf{H}, \gamma) = \frac{15^3}{16\sqrt{5}\pi^3} \frac{1}{\sigma_T^6(1-\gamma^2)^3} \exp\left[-\frac{3}{2\sigma_T^2(1-\gamma^2)}(2K_1^2 - 5K_2)\right] \quad (1)$$

where

$$K_1 = (T_{11} - \eta H_{11}) + (T_{22} - \eta H_{22}) + (T_{33} - \eta H_{33}) = k_1 - \eta h_1 \quad (2)$$

$$K_2 = (T_{11} - \eta H_{11})(T_{22} - \eta H_{22}) + (T_{11} - \eta H_{11})(T_{33} - \eta H_{33}) + (T_{22} - \eta H_{22})(T_{33} - \eta H_{33}) - (T_{12} - \eta H_{12})^2 - (T_{13} - \eta H_{13})^2 - (T_{23} - \eta H_{23})^2 = k_2 + \eta^2 h_2 - \eta h_1 k_1 + \eta \tau \quad (3)$$

$$\tau = T_{11}H_{11} + T_{22}H_{22} + T_{33}H_{33} + 2T_{12}H_{12} + 2T_{13}H_{13} + 2T_{23}H_{23} \quad (4)$$

$$k_1 = T_{11} + T_{22} + T_{33} \quad (5)$$

$$k_2 = T_{11}T_{22} + T_{11}T_{33} + T_{22}T_{33} - T_{12}^2 - T_{13}^2 - T_{23}^2 \quad (6)$$

$$h_1 = H_{11} + H_{22} + H_{33} \quad (7)$$

$$h_2 = H_{11}H_{22} + H_{11}H_{33} + H_{22}H_{33} - H_{12}^2 - H_{13}^2 - H_{23}^2 \quad (8)$$

$$\eta = \gamma\sigma_T/\sigma_H. \quad (9)$$

The corresponding unconditional marginal distributions  $p(\mathbf{T})$  and  $p(\mathbf{H})$  are multidimensional Gaussians, expressed using Doroshkevich’s formulas as

$$p(\mathbf{T}) = \frac{15^3}{16\sqrt{5}\pi^3} \frac{1}{\sigma_T^6} \exp\left[-\frac{3}{2\sigma_T^2}(2k_1^2 - 5k_2)\right], \quad p(\mathbf{H}) = \frac{15^3}{16\sqrt{5}\pi^3} \frac{1}{\sigma_H^6} \exp\left[-\frac{3}{2\sigma_H^2}(2h_1^2 - 5h_2)\right]. \quad (10)$$

Equation (1) generalizes Doroshkevich’s formulas (10) to include the fact that halos/voids may correspond to maxima/minima of the density field. Note in particular that  $p(\mathbf{T}|\mathbf{H}, \gamma)$  is a multivariate Gaussian distribution with mean  $b = \eta\mathbf{H}$  and covariance matrix  $(1-\gamma^2)\sigma_T^2\mathbf{A}/15$ , with  $\mathbf{A}$  given in Appendix A. Clearly, one can also consider the reverse distribution  $p(\mathbf{H}|\mathbf{T}, \gamma)$ , the expression of which is given in Rossi (2012). This is useful in order to make the connection and extend some results derived in Bardeen et al. (1986), the subject of a forthcoming publication.

It is also possible to express (1) in terms of the constrained eigenvalues of  $\mathbf{T}|\mathbf{H}$ , indicated with  $\zeta_i$  ( $i = 1, 2, 3$ ) and ordered as  $\zeta_1 \geq \zeta_2 \geq \zeta_3$ . The result is:

$$p(\zeta_1, \zeta_2, \zeta_3 | \gamma) = \frac{15^3}{8\sqrt{5}\pi} \frac{1}{\sigma_T^6 (1 - \gamma^2)^3} \exp \left[ -\frac{3}{2\sigma_T^2 (1 - \gamma^2)} (2K_1^2 - 5K_2) \right] (\zeta_1 - \zeta_2)(\zeta_1 - \zeta_3)(\zeta_2 - \zeta_3) \quad (11)$$

where in terms of constrained eigenvalues we now have:

$$K_1 = \zeta_1 + \zeta_2 + \zeta_3 = k_1 - \eta h_1 \quad (12)$$

$$K_2 = \zeta_1 \zeta_2 + \zeta_1 \zeta_3 + \zeta_2 \zeta_3 = k_2 + \eta^2 h_2 - \eta h_1 k_1 + \gamma \tau \quad (13)$$

$$\tau = \lambda_1 \xi_1 + \lambda_2 \xi_2 + \lambda_3 \xi_3 \quad (14)$$

$$k_1 = \lambda_1 + \lambda_2 + \lambda_3 \quad (15)$$

$$k_2 = \lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \lambda_2 \lambda_3 \quad (16)$$

$$h_1 = \xi_1 + \xi_2 + \xi_3 \quad (17)$$

$$h_2 = \xi_1 \xi_2 + \xi_1 \xi_3 + \xi_2 \xi_3 \quad (18)$$

$$\zeta_i = \lambda_i - \eta \xi_i. \quad (19)$$

The partial distributions  $p(\lambda_1, \lambda_2, \lambda_3)$  and  $p(\xi_1, \xi_2, \xi_3)$  are expressed by Doroshkevich's unconditional formulas as

$$p(\lambda_1, \lambda_2, \lambda_3) = \frac{15^3}{8\sqrt{5}\pi} \frac{1}{\sigma_T^6} \exp \left[ -\frac{3}{2\sigma_T^2} (2k_1^2 - 5k_2) \right] (\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)(\lambda_2 - \lambda_3) \quad (20)$$

$$p(\xi_1, \xi_2, \xi_3) = \frac{15^3}{8\sqrt{5}\pi} \frac{1}{\sigma_H^6} \exp \left[ -\frac{3}{2\sigma_H^2} (2h_1^2 - 5h_2) \right] (\xi_1 - \xi_2)(\xi_1 - \xi_3)(\xi_2 - \xi_3), \quad (21)$$

where  $\lambda_1, \lambda_2, \lambda_3$  are the eigenvalues of the shear tensor, while  $\xi_1, \xi_2, \xi_3$  are those of the density Hessian. Similarly, one can easily obtain formulas for the reverse probability functions. Note also that  $k_1 = \lambda_1 + \lambda_2 + \lambda_3$  is simply the overdensity  $\delta_T$  associated to the shear field, while  $\xi_1 + \xi_2 + \xi_3 = h_1 \equiv \delta_H$ . Later on, we will make use of the peak height  $\nu = \delta_T / \sigma_T$  and of the peak curvature  $x = \delta_H / \sigma_H$ .

Equations (1) and (11) allow one to develop a new algorithm to sample the constrained eigenvalues of the initial shear field, presented in the next section. It will be then straightforward to use this algorithm in order to implement an excursion set model for peaks and dips in Gaussian random fields.

### 3 THE PEAK/DIP EXCURSION-SET-BASED ALGORITHM

While the constrained eigenvalues of the initial shear field can be sampled directly from their probability distribution function (i.e. Equations 1 or 11), the new theoretical formalism developed by Rossi (2012) leads to a much simpler algorithm, which is particularly interesting because well-suited for the ellipsoidal collapse model (Section 5), and naturally integrable in the excursion set framework (Section 6). In addition, the algebra leading to the new computational procedure allows one to explain analytically how the halo (void) shape distributions are altered by the inclusion of the peak (dip) constraint, and offers a variety of applications and insights on the statistical description of the cosmic web (see Section 7 and Appendix B). In what follows, we present the mathematical aspects of the ‘peak/dip excursion-set-based’ algorithm, and describe in detail the novel computational procedure.

In particular, we are mainly interested in the distribution  $p(\mathbf{T}|\mathbf{H} > 0, \gamma)$ , the joint probability of observing a tidal field  $\mathbf{T}$  for the gravitational potential with a positive density curvature  $\mathbf{H} > 0$  (i.e. at density peak locations). Clearly,

$$p(\mathbf{T}|\mathbf{H} > 0, \gamma) = \frac{p(\mathbf{T}, \mathbf{H} > 0 | \gamma)}{p(\mathbf{H} > 0)} = \frac{\int_{\mathbf{H} > 0} p(\mathbf{H}) p(\mathbf{T}|\mathbf{H}, \gamma) d\mathbf{H}}{\int_{\mathbf{H} > 0} p(\mathbf{H}) d\mathbf{H}}, \quad (22)$$

where  $p(\mathbf{T}|\mathbf{H}, \gamma)$  and  $p(\mathbf{H})$  are given by Equations (1) and (10), and the integrals are 6-dimensional. A similar expression can be written in terms of the constrained distributions of eigenvalues, using Equation (11) instead. In principle, one should then compute the previous integral to obtain  $p(\mathbf{T}|\mathbf{H} > 0, \gamma)$ . However, it is easier to sample  $p(\mathbf{T}|\mathbf{H}, \gamma)$  and impose the condition  $\mathbf{H} > 0$  directly, so that we are effectively computing  $p(\mathbf{T}|\mathbf{H} > 0, \gamma)$ . Given the nature of  $p(\mathbf{T}|\mathbf{H}, \gamma)$  and  $p(\mathbf{H})$  expressed by (1) and (10), this can be readily achieved – as the following algebra will show.

In fact, because the elements of the density Hessian are drawn from a multivariate Gaussian distribution with zero mean and covariance matrix  $\sigma_H^2 \mathbf{A}/15$ , where  $\mathbf{A}$  is given by Equation (A8), one can simply obtain them by generating six independent

zero-mean unit-variance Gaussian random variates  $y_i$  ( $i = 1, 6$ ), and then determine the various components as:

$$\begin{aligned}
 H_{11} &= \frac{\sigma_H}{3} \left( -y_1 + \frac{2}{\sqrt{5}} y_2 \right) \\
 H_{22} &= \frac{\sigma_H}{3} \left( -y_1 - \frac{1}{\sqrt{5}} y_2 - \frac{3}{\sqrt{15}} y_3 \right) \\
 H_{33} &= \frac{\sigma_H}{3} \left( -y_1 - \frac{1}{\sqrt{5}} y_2 + \frac{3}{\sqrt{15}} y_3 \right) \\
 H_{12} &= H_{21} = \frac{\sigma_H}{\sqrt{15}} y_4 \\
 H_{13} &= H_{31} = \frac{\sigma_H}{\sqrt{15}} y_5 \\
 H_{23} &= H_{32} = \frac{\sigma_H}{\sqrt{15}} y_6.
 \end{aligned} \tag{23}$$

Similarly, the elements of the shear tensor are also drawn from a multivariate Gaussian distribution with mean zero and covariance matrix  $\sigma_T^2 \mathbf{A}/15$ . Hence, if  $z_i$  ( $i = 1, 6$ ) are other six independent zero-mean unit-variance Gaussian random variates, the components of the shear tensor are given by:

$$\begin{aligned}
 T_{11} &= \frac{\sigma_T}{3} \left( -z_1 + \frac{2}{\sqrt{5}} z_2 \right) \\
 T_{22} &= \frac{\sigma_T}{3} \left( -z_1 - \frac{1}{\sqrt{5}} z_2 - \frac{3}{\sqrt{15}} z_3 \right) \\
 T_{33} &= \frac{\sigma_T}{3} \left( -z_1 - \frac{1}{\sqrt{5}} z_2 + \frac{3}{\sqrt{15}} z_3 \right) \\
 T_{12} &= T_{21} = \frac{\sigma_T}{\sqrt{15}} z_4 \\
 T_{13} &= T_{31} = \frac{\sigma_T}{\sqrt{15}} z_5 \\
 T_{23} &= T_{32} = \frac{\sigma_T}{\sqrt{15}} z_6.
 \end{aligned} \tag{24}$$

Next, consider the 6-dimensional vector  $\mathbf{T}|\mathbf{H}$ , made from the components of the conditional shear field. Its elements are drawn from a multivariate Gaussian distribution  $p(\mathbf{T}|\mathbf{H}, \gamma)$  with mean  $b = \eta \mathbf{H}$  and covariance matrix  $(1 - \gamma^2) \sigma_T^2 \mathbf{A}/15$ , where the elements of the density Hessian are expressed by (23). Therefore, this implies that the components of  $\mathbf{T}|\mathbf{H}$  are distributed according to:

$$\begin{aligned}
 T_{11}|\mathbf{H}_{11} &= \eta H_{11} + \frac{\sigma_T \sqrt{1 - \gamma^2}}{3} \left( -l_1 + \frac{2}{\sqrt{5}} l_2 \right) \equiv \frac{\sigma_T}{3} \left( -m_1 + \frac{2}{\sqrt{5}} m_2 \right) \\
 T_{22}|\mathbf{H}_{22} &= \eta H_{22} + \frac{\sigma_T \sqrt{1 - \gamma^2}}{3} \left( -l_1 - \frac{1}{\sqrt{5}} l_2 - \frac{3}{\sqrt{15}} l_3 \right) \equiv \frac{\sigma_T}{3} \left( -m_1 - \frac{1}{\sqrt{5}} m_2 - \frac{3}{\sqrt{15}} m_3 \right) \\
 T_{33}|\mathbf{H}_{33} &= \eta H_{33} + \frac{\sigma_T \sqrt{1 - \gamma^2}}{3} \left( -l_1 - \frac{1}{\sqrt{5}} l_2 + \frac{3}{\sqrt{15}} l_3 \right) \equiv \frac{\sigma_T}{3} \left( -m_1 - \frac{1}{\sqrt{5}} m_2 + \frac{3}{\sqrt{15}} m_3 \right) \\
 T_{12}|\mathbf{H}_{12} &= T_{21}|\mathbf{H}_{21} = \eta H_{12} + \frac{\sigma_T \sqrt{1 - \gamma^2}}{\sqrt{15}} l_4 \equiv \frac{\sigma_T}{\sqrt{15}} m_4 \\
 T_{13}|\mathbf{H}_{13} &= T_{31}|\mathbf{H}_{31} = \eta H_{13} + \frac{\sigma_T \sqrt{1 - \gamma^2}}{\sqrt{15}} l_5 \equiv \frac{\sigma_T}{\sqrt{15}} m_5 \\
 T_{23}|\mathbf{H}_{23} &= T_{32}|\mathbf{H}_{32} = \eta H_{23} + \frac{\sigma_T \sqrt{1 - \gamma^2}}{\sqrt{15}} l_6 \equiv \frac{\sigma_T}{\sqrt{15}} m_6.
 \end{aligned} \tag{25}$$

In the previous expression, the various  $l_i$  ( $i = 1, 6$ ) are other six independent zero-mean unit-variance Gaussian variates, while the  $m_i$  ( $i = 1, 6$ ) are six independent Gaussian distributed variates with shifted mean  $\gamma y_i$  and reduced variance  $(1 - \gamma^2)$ , i.e.  $m_i = \gamma y_i + \sqrt{1 - \gamma^2} l_i$ .

Equations (23), (24) and (25) suggest a new excursion-set-based algorithm, in order to obtain the constrained eigenvalues of the matrix having components  $T_\alpha|\mathbf{H}_\alpha$  (see Appendix A for the definition of the index  $\alpha$ ) – supplemented by the condition of a positive curvature  $\mathbf{H}$  for the density field (or negative curvature, for voids). The procedure can be summarized as follows:

(i) Draw six independent zero-mean unit-variance Gaussian distributed variates  $y_i$  ( $i = 1, 6$ ), and determine the components  $H_\alpha$  of the density Hessian matrix via Equation (23). Compute the value of  $\sigma_H$  using (A6).

(ii) Calculate the eigenvalues  $\xi_1, \xi_2, \xi_3$  of the previous Hessian matrix, and check if they are all positive (negative). If so, proceed to the next step, otherwise try again. This will guarantee the Hessian to be positive (negative) definite (i.e. the Hessian is a real-symmetric matrix), which is the condition for maxima (minima) of the field. Note that this step is clearly not required if we want to sample only  $p(\mathbf{T}|\mathbf{H}, \gamma)$ .

(iii) Draw other six independent Gaussian distributed variates  $l_i$  ( $i = 1, 6$ ), with mean zero and variance one, and determine the  $T_\alpha|H_\alpha$  components via Equation (25), using the previously accepted values  $H_\alpha$  from (23) – while  $\sigma_T$  is determined via (A6). Since we require the density Hessian to be positive definite, this means that we are effectively sampling the conditional probability  $p(\mathbf{T}|\mathbf{H} > 0, \gamma)$  – or  $p(\mathbf{T}|\mathbf{H} < 0, \gamma)$  for a negative definite Hessian.

(iv) Calculate and store the constrained eigenvalues  $\zeta_i$  of the matrix having components  $T_\alpha|H_\alpha > 0$ .

We leave in Appendix B some more insights on the main conditional formula (1), which readily follow from the previous Equations (23), (24) and (25). In the next sections, we will show a few applications of the new algorithm – with particular emphasis on the conditional distributions of shape parameters in triaxial models of structure formation (i.e. ellipticity and prolateness). Additional applications will be presented in forthcoming publications.

#### 4 CONDITIONAL DISTRIBUTIONS AND PROBABILITIES: NUMERICAL TESTS

The algorithm illustrated in the previous section allows one to test and confirm several formulas derived by Rossi (2012), regarding conditional distributions and probabilities subjected to the extremum constraint. We present here results of the comparison (i.e. theory versus numerical), where for simplicity we set  $\sigma_T = \sigma_H \equiv 1$ ; this corresponds to adopt ‘reduced variables’, as done in Rossi (2012). Note that we show explicitly the dependencies on  $\sigma$  values in the following expressions, although we set them to unity in the mock tests.

##### 4.1 Individual conditional distributions

The individual conditional distributions of shear eigenvalues, for a given correlation strength  $\gamma$  with the density field (which encapsulate the peak/dip constraint), are given by (Rossi 2012):

$$p(\zeta_1|\gamma) = \frac{\sqrt{5}}{12\pi\sigma_T} \left[ \frac{20}{(1-\gamma^2)} \frac{\zeta_1}{\sigma_T} \exp\left(-\frac{9}{2(1-\gamma^2)} \frac{\zeta_1^2}{\sigma_T^2}\right) - \frac{\sqrt{2\pi}}{(1-\gamma^2)^{3/2}} \exp\left(-\frac{5}{2(1-\gamma^2)} \frac{\zeta_1^2}{\sigma_T^2}\right) \times \right. \\ \left. \operatorname{erfc}\left(-\frac{\sqrt{2}}{\sqrt{1-\gamma^2}} \frac{\zeta_1}{\sigma_T}\right) \left[(1-\gamma^2) - 20 \frac{\zeta_1^2}{\sigma_T^2}\right] + \frac{3\sqrt{3\pi}}{\sqrt{1-\gamma^2}} \exp\left(-\frac{15}{4(1-\gamma^2)} \frac{\zeta_1^2}{\sigma_T^2}\right) \operatorname{erfc}\left(-\frac{\sqrt{3}}{2\sqrt{1-\gamma^2}} \frac{\zeta_1}{\sigma_T}\right) \right] \quad (26)$$

$$p(\zeta_2|\gamma) = \frac{\sqrt{15}}{2\sqrt{\pi}} \frac{1}{\sigma_T \sqrt{1-\gamma^2}} \exp\left[-\frac{15}{4(1-\gamma^2)} \frac{\zeta_2^2}{\sigma_T^2}\right] \quad (27)$$

$$p(\zeta_3|\gamma) = -\frac{\sqrt{5}}{12\pi\sigma_T} \left[ \frac{20}{(1-\gamma^2)} \frac{\zeta_3}{\sigma_T} \exp\left(-\frac{9}{2(1-\gamma^2)} \frac{\zeta_3^2}{\sigma_T^2}\right) + \frac{\sqrt{2\pi}}{(1-\gamma^2)^{3/2}} \exp\left(-\frac{5}{2(1-\gamma^2)} \frac{\zeta_3^2}{\sigma_T^2}\right) \times \right. \\ \left. \operatorname{erfc}\left(\frac{\sqrt{2}}{\sqrt{1-\gamma^2}} \frac{\zeta_3}{\sigma_T}\right) \left[(1-\gamma^2) - 20 \frac{\zeta_3^2}{\sigma_T^2}\right] - \frac{3\sqrt{3\pi}}{\sqrt{1-\gamma^2}} \exp\left(-\frac{15}{4(1-\gamma^2)} \frac{\zeta_3^2}{\sigma_T^2}\right) \operatorname{erfc}\left(\frac{\sqrt{3}}{2\sqrt{1-\gamma^2}} \frac{\zeta_3}{\sigma_T}\right) \right]. \quad (28)$$

The previous formulas show explicitly that the conditional distributions of shear eigenvalues are Doroshkevich-like expressions, with shifted mean and reduced variance. Different panels in Figure 1 display plots of these distributions, contrasted with results from the algorithm presented in Section 3, where 500,000 mock realizations are considered. In particular, note the symmetry between  $p(\zeta_1|\gamma)$  and  $p(\zeta_3|\gamma)$ . We expect the mean values and variances of those distributions to be:

$$\langle \zeta_1|\gamma \rangle = \frac{3}{\sqrt{10\pi}} \sigma_T \sqrt{1-\gamma^2}, \quad \sigma_{\zeta_1|\gamma}^2 = \frac{13\pi - 27}{30\pi} \sigma_T^2 (1-\gamma^2) \quad (29)$$

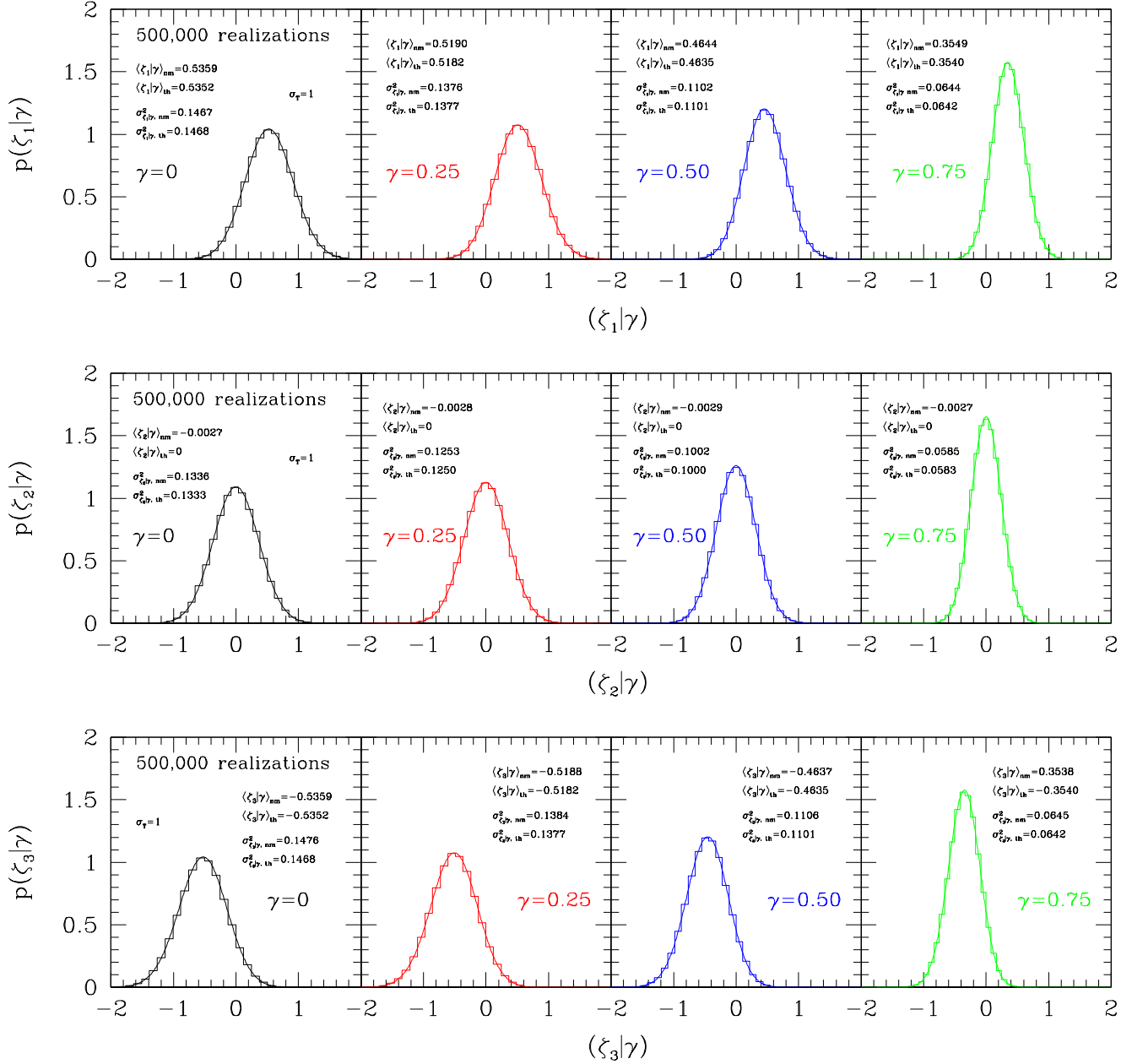
$$\langle \zeta_2|\gamma \rangle = 0, \quad \sigma_{\zeta_2|\gamma}^2 = \frac{2}{15} \sigma_T^2 (1-\gamma^2) \quad (30)$$

$$\langle \zeta_3|\gamma \rangle \equiv -\langle \zeta_1|\gamma \rangle = -\frac{3}{\sqrt{10\pi}} \sigma_T \sqrt{1-\gamma^2}, \quad \sigma_{\zeta_3|\gamma}^2 \equiv \sigma_{\zeta_1|\gamma}^2 = \frac{13\pi - 27}{30\pi} \sigma_T^2 (1-\gamma^2). \quad (31)$$

In the various panels, both theoretical and numerical expectations are reported and found to be in excellent agreement. For example, when  $\gamma = 0.50$ ,  $\langle \zeta_1|\gamma \rangle_{\text{num}} = 0.4644$  while the expected theoretical value is  $\langle \zeta_1|\gamma \rangle_{\text{th}} = 0.4635$ , and for the variances  $\sigma_{\zeta_1|\gamma}^{2,\text{num}} = 0.1102$  where the theoretical expectation is  $\sigma_{\zeta_1|\gamma}^{2,\text{th}} = 0.1101$ .

Obviously, to see explicitly the effect of the peak constraint, one needs to make cuts in  $\xi_i$  from the partial conditional distributions  $\lambda_i|\xi_i$ , since  $\zeta_i = \lambda_i - \eta\xi_i$ . In particular, expressions for  $p(\lambda_i|\mathbf{H} > 0, \gamma)$  can be derived from (26–28) via substitution of variables and integration over  $\mathbf{H} > 0$ . For example, it is straightforward to obtain an analytic formula for  $p(\lambda_3|\mathbf{H} > 0, \gamma)$ . In Section 5, we will return on these issues in more detail, in connection with the conditional ellipticity and prolateness for dark matter halos.

Another important distribution is  $p(\Delta_{T|\mathbf{H}}|\zeta_3 > 0, \gamma)$ , with  $\Delta_{T|\mathbf{H}} \equiv K_1 = \zeta_1 + \zeta_2 + \zeta_3$  being the sum of the constrained



**Figure 1.** Individual conditional probabilities  $p(\xi_1|\gamma)$  [top],  $p(\xi_2|\gamma)$  [middle], and  $p(\xi_3|\gamma)$  [bottom] of the initial shear field in the peak/dip picture. Solid curves are Equations (26), (27), (28), while histograms are obtained from 500,000 mock realizations via the algorithm described in Section 3. Each panel displays the numerical and theoretical expectations for the mean values and variances of these distributions; their agreement is excellent.

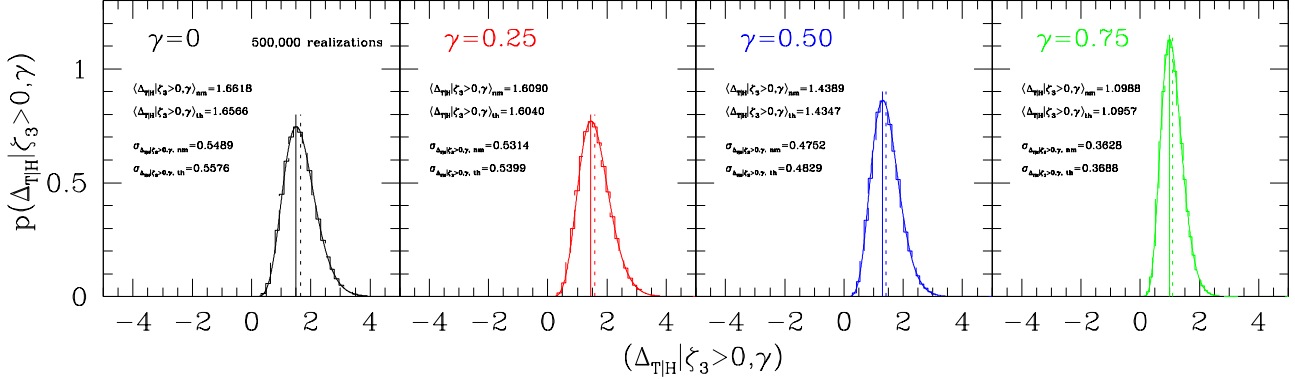
eigenvalues, namely the probability distribution of  $\Delta_{\text{T|H}}$  confined in the regions with  $\xi_3 > 0$  (i.e. all positive eigenvalues):

$$\begin{aligned}
 p(\Delta_{\text{T|H}}|\xi_3 > 0, \gamma) &= -\frac{75\sqrt{5}}{8\pi\sigma_{\text{T}}^2} \frac{\Delta_{\text{T|H}}}{(1-\gamma^2)} \exp\left(-\frac{9}{8} \frac{\Delta_{\text{T|H}}^2}{\sigma_{\text{T}}^2(1-\gamma^2)}\right) + \\
 &+ \frac{25}{4\sqrt{2\pi}} \frac{1}{\sigma_{\text{T}}\sqrt{1-\gamma^2}} \exp\left(-\frac{\Delta_{\text{T|H}}^2}{2\sigma_{\text{T}}^2(1-\gamma^2)}\right) \left[ \text{erf}\left(\frac{\sqrt{10}\Delta_{\text{T|H}}}{4\sigma_{\text{T}}\sqrt{1-\gamma^2}}\right) + \text{erf}\left(\frac{\sqrt{10}\Delta_{\text{T|H}}}{2\sigma_{\text{T}}\sqrt{1-\gamma^2}}\right) \right].
 \end{aligned} \tag{32}$$

This distribution is expected to have mean value

$$\langle \Delta_{\text{T|H}}|\xi_3 > 0, \gamma \rangle = \frac{25\sqrt{10}}{144\sqrt{\pi}} (3\sqrt{6} - 2) \sigma_{\text{T}}\sqrt{1-\gamma^2} \simeq 1.66 \sigma_{\text{T}}\sqrt{1-\gamma^2}. \tag{33}$$

It is also easy to see that the maximum of  $p(\Delta_{\text{T|H}}|\xi_3 > 0, \gamma)$  is reached when  $\Delta_{\text{T|H}} \simeq 1.5 \sigma_{\text{T}}\sqrt{1-\gamma^2}$ , and with a more



**Figure 2.** Conditional probability distribution  $p(\Delta_{T|H} | \zeta_3 > 0, \gamma)$ . Solid lines are obtained from Equation (32), for different values of the correlation parameter  $\gamma$ , while histograms are drawn from 500,000 realizations via the algorithm described in Section 3. Note the excellent agreement between numerical results and theoretical predictions for the corresponding mean and rms values, reported in the various panels. In addition, solid vertical lines show the expected maximum value of each distribution, while dotted lines represent their corresponding mean value.

elaborate calculation its variance can be estimated analytically. The result is:

$$\sigma_{\Delta_{T|H} | \zeta_3 > 0, \gamma}^2 = \frac{25}{4\pi} \sigma_T^2 (1 - \gamma^2) \left[ \arctan(\sqrt{5}/2) + \arctan(\sqrt{5}) - \frac{11\sqrt{5}}{54} - \frac{125}{2592} (3\sqrt{6} - 2)^2 \right] \simeq 0.31 \sigma_T^2 (1 - \gamma^2). \quad (34)$$

Figure 2 confirms the previous relations, by contrasting Equations (32), (33) and (34) with numerical results from 500,000 realizations from the algorithm presented in Section 3. We find good agreement, and recover correctly the expected mean values and variances of the distributions. For example, when  $\gamma = 0.25$ , we find  $\langle \Delta_{T|H} | \zeta_3 > 0, \gamma \rangle_{\text{num}} = 1.6090$  while the expected theoretical value is  $\langle \Delta_{T|H} | \zeta_3 > 0, \gamma \rangle_{\text{th}} = 1.6040$ , and for the rms values we find  $\sigma_{\Delta_{T|H} | \zeta_3 > 0, \gamma}^{\text{num}} = 0.5314$  where the expected theoretical value is  $\sigma_{\Delta_{T|H} | \zeta_3 > 0, \gamma}^{\text{th}} = 0.5399$ . In the absence of correlations between the potential and density fields (i.e. when  $\gamma = 0$ ), all the previous expressions reduce consistently to the unconditional limit of Lee & Shandarin (1998).

## 4.2 Distribution of peak heights

An important quantity which plays a major role in peaks theory (Bardeen et al. 1986) is the distribution of peak heights, namely  $p(\nu | \mathbf{H} > 0, \gamma)$ , where  $\nu = \delta_T / \sigma_T$  and the overdensity  $\delta_T$  is the trace of the shear tensor – defined in Section 2. One can easily obtain a simplified analytic expression for this distribution as follows. First, normalize Equation (12) by  $\sigma_T$ , defining  $N = \Delta_{T|H} / \sigma_T$  and  $x = \delta_H / \sigma_H$ , so that  $N = \nu - \gamma x$ . In this framework,  $x$  is the peak curvature of Bardeen et al. (1986) if Gaussian filters are used for computing the moments of the smoothed power spectrum (Equation A6). Since  $p(N | \gamma)$  is simply a Gaussian with mean zero and variance  $(1 - \gamma^2)$  – see Rossi (2012) – and because of Equation (12), then clearly  $p(\nu | x, \gamma)$  is also Gaussian, with mean  $\gamma x$  and reduced variance  $(1 - \gamma^2)$ . Note in fact that  $p(\nu, x | \gamma)$  is a bivariate Gaussian, because both  $\nu$  and  $x$  are normally distributed random variates. This implies that  $\langle \nu | x, \gamma \rangle = \gamma x$ , and clearly  $\langle \nu | \mathbf{H} > 0, \gamma \rangle = \gamma \langle x | \zeta_3 > 0 \rangle \equiv \gamma \langle \nu | \lambda_3 > 0 \rangle$  (Bardeen et al. 1986) – where  $\langle \nu | \lambda_3 > 0 \rangle = 1.6566$  according to Equation (21) of Lee & Shandarin (1998). Starting from a bivariate Gaussian with the previous expected mean value, it is then direct to derive a fairly good approximation for the distribution of peak heights, namely

$$p(\nu | \mathbf{H} > 0, \gamma) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(\nu - \chi)^2}{2}\right] \left[1 + \text{erf}\left(\frac{\gamma(\nu - \chi)}{\sqrt{2(1 - \gamma^2)}}\right)\right] \quad (35)$$

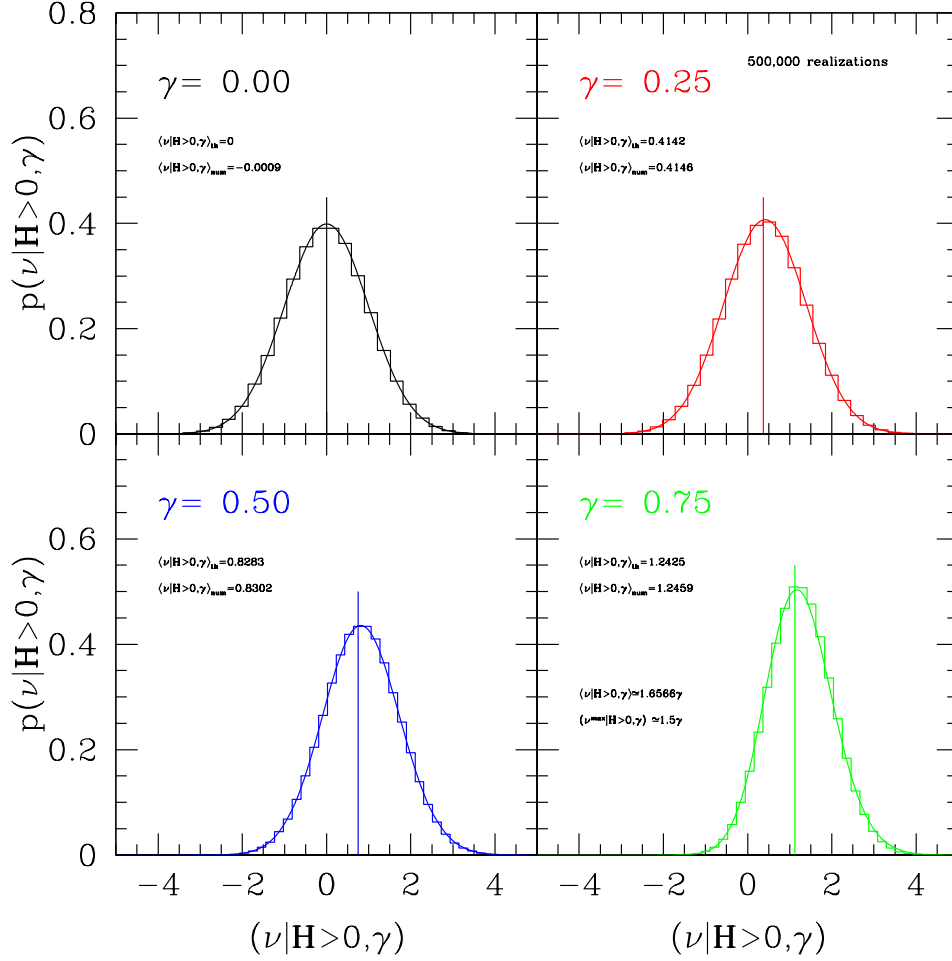
where

$$\chi = \frac{\sqrt{2}}{\sqrt{\pi}} \left[ \frac{25\sqrt{5}}{144} (3\sqrt{6} - 2) - 1 \right] \gamma \simeq 0.86\gamma. \quad (36)$$

In essence,  $p(\nu | \mathbf{H} > 0, \gamma)$  is a Gaussian with shifted mean modulated by the role of the peak curvature, and it consistently reduces to a zero-mean unit-variance Gaussian distribution when  $\gamma = 0$  – as expected. Additionally, the previous distribution can be shown to be equivalent to  $p(\nu) [\int_0^\infty p(\xi_3 | \gamma, \nu) d\xi_3 / p(\xi_3 > 0)]$ , where the term in square brackets quantifies the effect of the peak constrain – and so the scale at which the difference between peaks and random positions would appear.

Figure 3 shows plots of  $p(\nu | \mathbf{H} > 0, \gamma)$  for different values of  $\gamma$ . In the various panels, vertical lines are the expected mean values; we find good agreement between numerical estimates and theoretical expectations. For example, when  $\gamma = 0.25$ ,  $\langle \nu | \mathbf{H} > 0, \gamma \rangle_{\text{num}} = 0.4146$  while  $\langle \nu | \mathbf{H} > 0, \gamma \rangle_{\text{th}} = 0.4142$ . Solid curves in the figure are drawn from Equation (35), an approximation which is particularly good for lower values of the correlation strength.





**Figure 3.** Distribution of peak heights,  $p(\nu|H > 0, \gamma)$ , for different values of the correlation parameter  $\gamma$  – as specified in the various panels. The numerical values for the mean of the distribution as a function of  $\gamma$  are indicated in each panel, and also marked with vertical lines; they are in good agreement with theoretical expectations. Solid curves are drawn from Equation (35).

## 5 CONDITIONAL ELLIPTICITY AND PROLATENESS

The theoretical framework outlined in Section 2, along with the algebra leading to the new computational procedure presented in Section 3, provides a natural way to include the peak/dip constraint into the initial shape distributions of halos and voids. This is a key aspect for implementing an excursion set algorithm for peaks and dips in Gaussian random fields (which will be discussed in Section 6), and has direct applications in the context of triaxial models of structure formation. In particular, our description combines the formalism of Bardeen et al. (1986) – based on the density field – with that of Bond & Myers (1996) – focused on the shear field.

In what follows, first we introduce some useful definitions regarding shape parameters (ellipticity and prolateness); we then provide a new expression for the joint conditional distribution of halo shapes given the density peak constraint, derived from Equation (11), which generalizes some previous literature work. The formula has important implications for the modeling of non-spherical dark matter halos, in relation to their initial shape distribution. Finally, we briefly discuss how the shear ellipticity and prolateness (and so the initial halo shapes) are modified by the inclusion of the density peak condition.

Our analysis is based on the constrained eigenvalues of the initial shear field (Equations 11–19), and we use the new algorithm to support our description. While here we only consider the case for dark matter halos, it is straightforward to deal with voids. In a forthcoming companion publication, we will present a more detailed study on the morphology of both halos and voids in the peak/dip picture, where we also investigate the modifications induced by primordial non-Gaussianity on their shapes (see also Section 7).

### 5.1 Shape parameters: general definitions

In triaxial models of halo formation, such as the ellipsoidal collapse (Icke 1973; White & Silk 1979; Barrow & Silk 198; Kuhlman et al. 1996; Bond & Myers 1996), it is customary to characterize the shape of a region by its ellipticity and prolateness. In

particular, in the ‘peak-patch’ approach of Bond & Myers (1996), the shape parameters are associated with the eigenvalues of the shear field ( $\lambda_i$  in our notation). Considering the mapping ( $\lambda_1, \lambda_2, \lambda_3 \rightarrow e_T, p_T, \delta_T$ ), where the subscript  $T$  refers to the tidal tensor, we can write:

$$\delta_T \equiv k_1 = \lambda_1 + \lambda_2 + \lambda_3, \quad e_T = \frac{\lambda_1 - \lambda_3}{2\delta_T}, \quad p_T = \frac{\lambda_1 + \lambda_3 - 2\lambda_2}{2\delta_T} \equiv e_T - \frac{\lambda_2 - \lambda_3}{\delta_T}, \quad (37)$$

where  $e_T$  and  $p_T$  are the ‘unconditional’ shear ellipticity and prolateness. In particular, if  $\delta_T > 0$  then  $e_T > 0$  and therefore  $-e_T \leq p_T \leq e_T$ .

In the ‘peaks theory’ of Bardeen et al. (1986), instead, ellipticity ( $e_H$ ) and prolateness ( $p_H$ ) are associated with the eigenvalues of the density field ( $\xi_i$  in our notation). Hence, given the mapping ( $\xi_1, \xi_2, \xi_3 \rightarrow e_H, p_H, \delta_H$ ) where now the subscript  $H$  refers to the Hessian of the density field, one has:

$$\delta_H \equiv h_1 = \xi_1 + \xi_2 + \xi_3, \quad e_H = \frac{\xi_1 - \xi_3}{2\delta_H}, \quad p_H = \frac{\xi_1 + \xi_3 - 2\xi_2}{2\delta_H} \equiv e_H - \frac{\xi_2 - \xi_3}{\delta_H}. \quad (38)$$

As in the previous case, if  $\delta_H > 0$  then  $e_H > 0$  and  $-e_H \leq p_H \leq e_H$ .

It is therefore natural to consider the ‘conditional’ mapping ( $\zeta_1, \zeta_2, \zeta_3 \rightarrow E_{T|H}, P_{T|H}, \Delta_{T|H}$ ) and define:

$$\Delta_{T|H} \equiv K_1 = \zeta_1 + \zeta_2 + \zeta_3, \quad E_{T|H} = \frac{\zeta_1 - \zeta_3}{2\Delta_{T|H}}, \quad P_{T|H} = \frac{\zeta_1 + \zeta_3 - 2\zeta_2}{2\Delta_{T|H}} \equiv E_{T|H} - \frac{\zeta_2 - \zeta_3}{\Delta_{T|H}}, \quad (39)$$

where  $\zeta_i$  are the constrained shear eigenvalues – given by Equation (19). In what follows, we will refer to  $E_{T|H}$  and  $P_{T|H}$  as to the conditional shear ellipticity and prolateness, respectively; we will also relate these quantities to the unconditional expressions previously introduced, involving ( $e_T, p_T$ ) and ( $e_H, p_H$ ).

## 5.2 Joint conditional distribution of shear ellipticity and prolateness in the peak/dip picture

Combining (37) with Doroshkevich’s formula (20), it is straightforward to derive the ‘unconditional’ distribution of  $e_T$  and  $p_T$  given  $\delta_T$ ,  $g(e_T, p_T | \delta_T)$ , related to the gravitational potential. In particular,

$$g(\delta_T, e_T, p_T) = p(\delta_T)g(e_T, p_T | \delta_T) \quad (40)$$

where  $p(\delta_T)$  is simply a Gaussian with zero mean and variance  $\sigma_T^2$ , and

$$g(e_T, p_T | \delta_T) = \frac{15^3}{3\sqrt{10\pi}} \left( \frac{\delta_T}{\sigma_T} \right)^5 e_T (e_T^2 - p_T^2) \exp \left[ -\frac{5}{2} \left( \frac{\delta_T}{\sigma_T} \right)^2 (3e_T^2 + p_T^2) \right]. \quad (41)$$

Equation (40) implies that the unconditional joint distribution of shear ellipticity and prolateness is independent of that of the overdensity  $\delta_T$ . Similarly, inserting (38) in Doroshkevich’s formula (21), one obtains

$$g(\delta_H, e_H, p_H) = p(\delta_H)g(e_H, p_H | \delta_H) \quad (42)$$

for the quantities related to the density Hessian, where  $p(\delta_H)$  is a Gaussian with zero mean and variance  $\sigma_H^2$ , while

$$g(e_H, p_H | \delta_H) = \frac{15^3}{3\sqrt{10\pi}} \left( \frac{\delta_H}{\sigma_H} \right)^5 e_H (e_H^2 - p_H^2) \exp \left[ -\frac{5}{2} \left( \frac{\delta_H}{\sigma_H} \right)^2 (3e_H^2 + p_H^2) \right]. \quad (43)$$

Along the same lines, combining (39) with (11) and using the definitions introduced in the previous section, it is direct to obtain

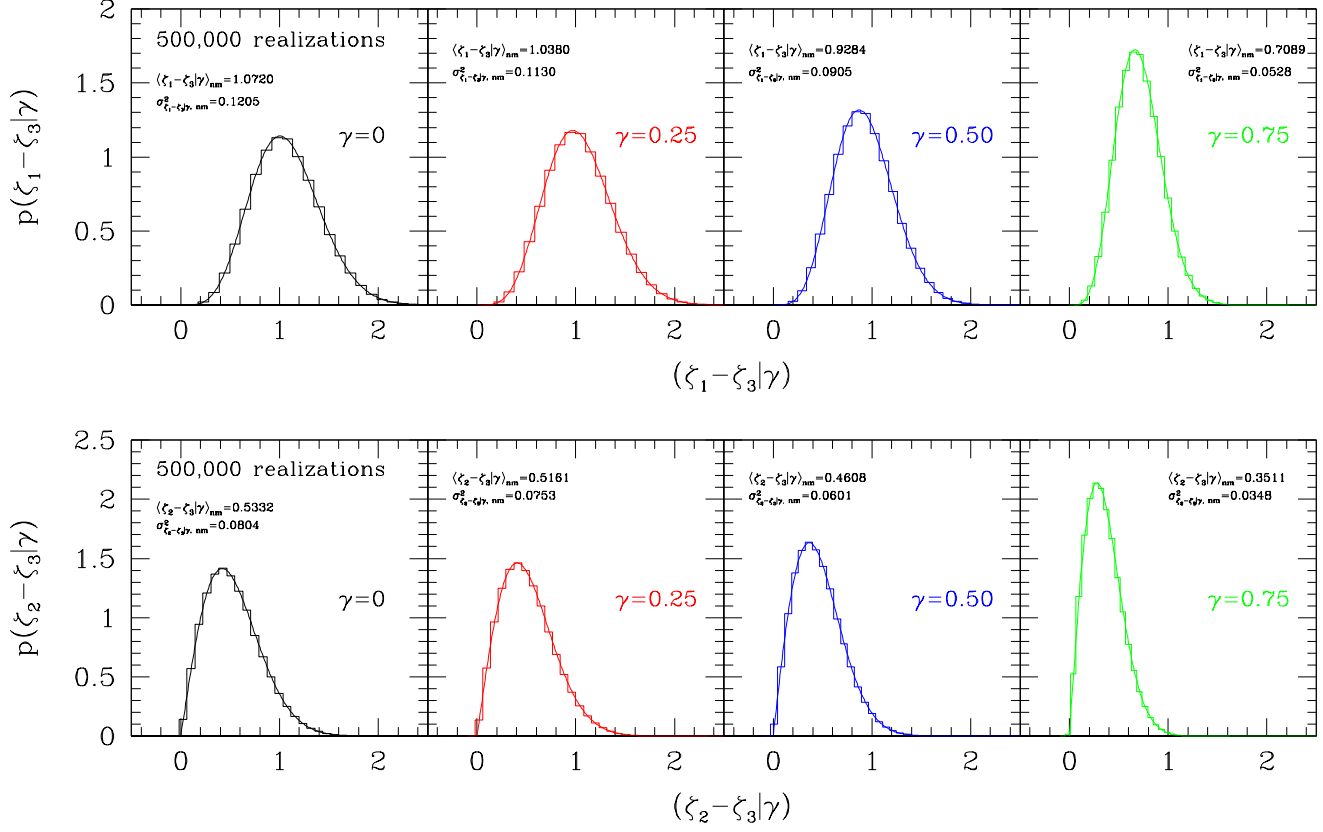
$$G(\Delta_{T|H}, E_{T|H}, P_{T|H} | \gamma) = p(\Delta_{T|H} | \gamma) G(E_{T|H}, P_{T|H} | \Delta_{T|H}, \gamma) \quad (44)$$

where  $p(\Delta_{T|H} | \gamma)$  is a Gaussian distribution with mean zero and variance  $\sigma_T^2(1 - \gamma^2)$  – i.e. Equation (57) in Rossi (2012) – and

$$G(E_{T|H}, P_{T|H} | \Delta_{T|H}, \gamma) = \frac{15^3}{3\sqrt{10\pi}} \left( \frac{\Delta_{T|H}}{\sigma_T} \right)^5 \frac{1}{(1 - \gamma^2)^{5/2}} E_{T|H} (E_{T|H}^2 - P_{T|H}^2) \exp \left[ \frac{5}{2(1 - \gamma^2)} \left( \frac{\Delta_{T|H}}{\sigma_T} \right)^2 (3E_{T|H}^2 + P_{T|H}^2) \right]. \quad (45)$$

The previous expression (45) is the joint conditional distribution of shear ellipticity and prolateness, given  $\Delta_{T|H}$  and a correlation strength  $\gamma$  with the density field. Clearly,  $G(E_{T|H}, P_{T|H} | \Delta_{T|H}, \gamma)$  is also independent of the distribution of  $(\Delta_{T|H} | \gamma)$ . Equation (45) generalizes the standard joint distribution of halo shape parameters to include the peak constraint, and is another main result of this paper. Note that, in the absence of correlation (i.e. when  $\gamma = 0$ ),  $E_{T|H} \rightarrow e_T$ ,  $P_{T|H} \rightarrow p_T$ ,  $\Delta_{T|H} \rightarrow \delta_T$ , so that (45) reduces consistently to the ‘unconditional’ Doroshkevich’s limit (41).

From this joint conditional probability function, it is possible to derive the partial conditional distributions for the shear ellipticity and prolateness given the peak constrain (which we will discuss next), and their expressions at density peak locations, when the condition  $\mathbf{H} > 0$  is satisfied.



**Figure 4.** [Top] Distribution of  $(\zeta_1 - \zeta_3 | \gamma)$ , for different values of the correlation strength  $\gamma$  – as indicated in the panels. Its shape is directly related to the shear conditional ellipticity in the peak/dip picture – see Equation (46). [Bottom] Distribution of  $(\zeta_2 - \zeta_3 | \gamma)$ , for different values of the correlation strength  $\gamma$ . This distribution is instead related to the shear conditional prolateness in the peak/dip picture – see Equation (47). The numerical values for the mean and variance of the distributions, as a function of  $\gamma$ , are also indicated in the figures. In both cases, solid curves are obtained via numerical integrations starting from the joint conditional distribution of eigenvalues (11). Histograms are drawn from 500,000 realizations, using the algorithm described in Section 3.

### 5.3 Initial distributions of triaxial halo shapes at density peak locations

We can readily express the conditional distributions of shear ellipticity and prolateness in the peak/dip picture in terms of the unconditional quantities (37) and (38). This is done simply by recalling that the constrained shear eigenvalues are given by  $\zeta_i = \lambda_i - \eta \xi_i$ , according to Equation (19).

It is then direct to obtain:

$$E_{T|H} \Delta_{T|H} = \frac{(\zeta_1 - \zeta_3)}{2} \equiv \frac{(\lambda_1 - \lambda_3)}{2} - \eta \frac{(\xi_1 - \xi_3)}{2} = \delta_T e_T - \eta \delta_H e_H. \quad (46)$$

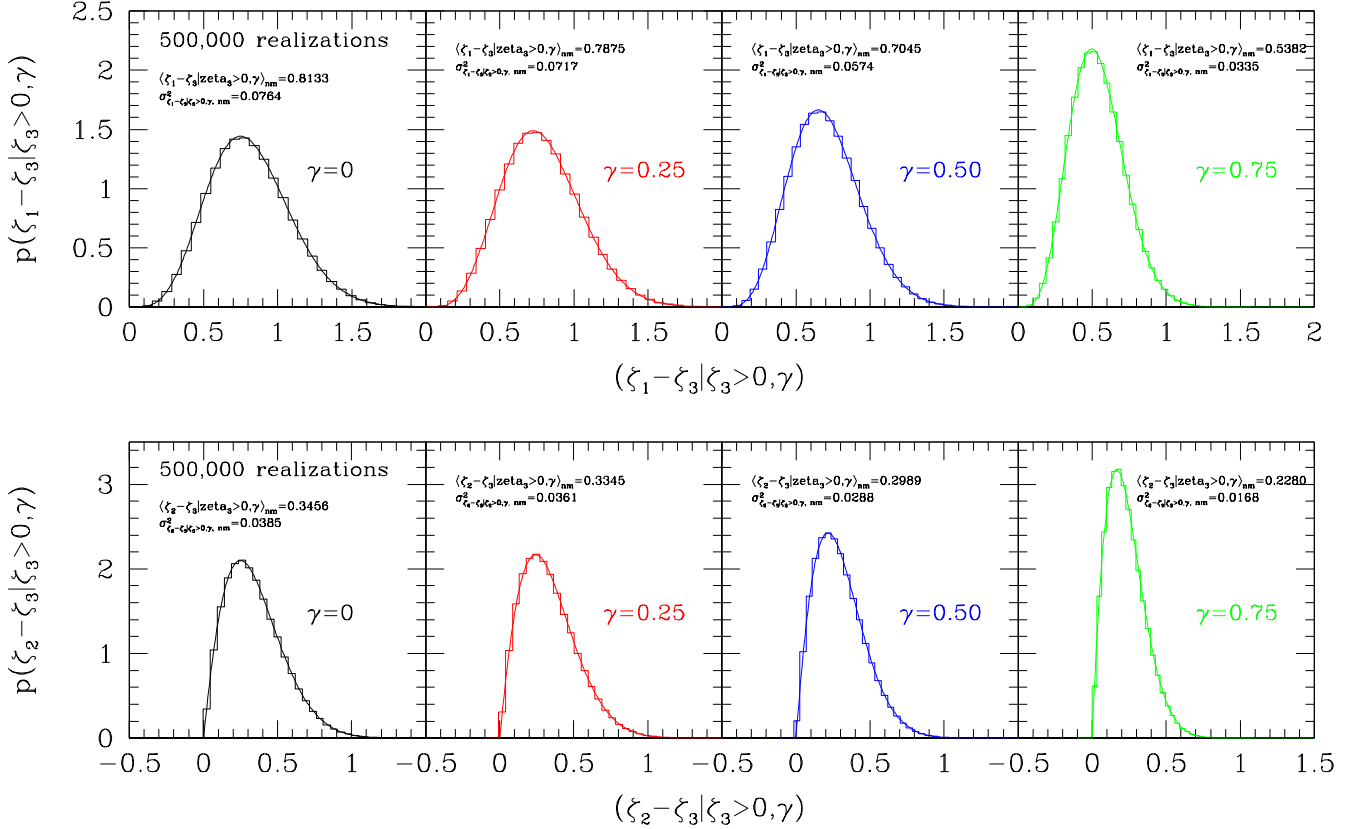
The previous relation implies that, for a given  $\Delta_{T|H}$ , the conditional shear ellipticity will have its mean value shifted by the presence of the peak constraint; the entity of the shift has to be ascribed to the additional factor  $\eta \delta_H e_H$ , which is given by the density ellipticity (essentially, the latter term quantifies the role of the peak curvature). The top panel of Figure 4 shows the distribution of  $(\zeta_1 - \zeta_3 | \gamma)$  for different values of the correlation strength  $\gamma$ , namely the combination of constrained shear eigenvalues which controls the conditional ellipticity – according to Equation (46). Histograms are drawn from 500,000 realizations using the algorithm described in Section 3, while solid curves are obtained via numerical integration – starting from the joint conditional distribution of eigenvalues (11). In the various panels, we also provide the numerical values for the mean and variance of the distribution, as a function of  $\gamma$ .

Similarly,

$$P_{T|H} \Delta_{T|H} = \frac{(\zeta_1 + \zeta_3 - 2\zeta_2)}{2} = \frac{(\lambda_1 + \lambda_3 - 2\lambda_2)}{2} - \eta \frac{(\xi_1 + \xi_3 - 2\xi_2)}{2} = \delta_T p_T - \eta \delta_H p_H \equiv E_{T|H} \Delta_{T|H} - (\zeta_2 - \zeta_3). \quad (47)$$

This expression shows that the conditional shear prolateness can be obtained from the conditional shear ellipticity (46) and the combination  $(\zeta_2 - \zeta_3)$  of the constrained shear eigenvalues. The bottom panel of Figure 4 displays the distribution of  $(\zeta_2 - \zeta_3 | \gamma)$ , for different values of the correlation strength  $\gamma$ . Again, solid curves are derived from a numerical integration of Equation (11).

Of particular interest are also the distributions  $g(e_{T|H} > 0, \gamma)$  and  $g(p_{T|H} > 0, \gamma)$ , which will provide the initial triaxial



**Figure 5.** Distributions  $p(\xi_1 - \xi_3 | \xi_3 > 0, \gamma)$  [top] and  $p(\xi_2 - \xi_3 | \xi_3 > 0, \gamma)$  [bottom] for different values of the correlation strength  $\gamma$ . In both figures, histograms are drawn from 500,000 realizations using the algorithm described in Section 3, while solid curves are obtained via numerical integrations of the joint conditional distribution of eigenvalues (11). In the various panels, we also provide the numerical values for the mean and variance of the distributions, as a function of  $\gamma$ . In particular, their mean values at  $\gamma = 0$  allow one to quantify the shift in the mean of the shear ellipticity and prolateness caused by the peak constraint.

shapes of dark matter halos at peak locations. They are related to the previous expressions, and can be readily obtained within the outlined formalism; clearly, they will reduce to the standard (or unconditional) shear ellipticity and prolateness when  $\gamma = 0$ , in the absence of correlation. We will present a dedicated study focused on the morphology of halos and voids in the peak/dip picture, where we characterize in detail those distributions, especially in relation to the ellipsoidal collapse model. We anticipate here that several analytic and insightful results on their shapes can be derived, starting from (45) and using (11), (46) and (47). For example, according to (46), when  $\mathbf{H} > 0$  (which is equivalent to impose the condition  $\xi_3 > 0$  on an ordered set of density Hessian eigenvalues), then the quantity  $\langle \xi_1 - \xi_3 | \xi_3 > 0 \rangle$  will essentially provide by how much the mean value of the shear ellipticity has been shifted by the peak constrain. For the prolateness the situation is slightly more complicated, but according to (47) the shift can be derived by the additional knowledge of  $\langle \xi_2 - \xi_3 | \xi_3 > 0 \rangle$ . To this end, Figure 5 shows the distributions of  $(\xi_1 - \xi_3 | \xi_3 > 0, \gamma)$  and  $(\xi_2 - \xi_3 | \xi_3 > 0, \gamma)$ , for different values of the correlation strength  $\gamma$  (note that the quantities just discussed are their mean values when  $\gamma = 0$ ). As in the previous plot, histograms are drawn from 500,000 realizations using the algorithm described in Section 3, while solid curves are obtained via numerical integrations – starting from the joint conditional distribution of eigenvalues (11). In the various panels, we also provide the numerical values for the mean and variance of the distribution, as a function of  $\gamma$ .

As mentioned before, we will return in more detail on these distributions in a companion publication, to quantify the impact of the peak constrain on their shapes. We will also make the connection with the work of Bardeen et al. (1986) more explicit – see their Appendix C. It will be also interesting to include the role of the peculiar gravity field in our description, along the lines of van de Weygaert & Bertschinger (1996), as well as to extend the work of Desjacques (2008) on the joint statistics of the shear tensor and on the dynamical aspect of the environmental dependence, within this formalism.

## 6 EXCURSION SET APPROACH FOR PEAKS AND DIPS IN GAUSSIAN RANDOM FIELDS

The ability to distinguish between random positions and peaks/dips, contained in Equations (1) or (11) and achieved by the algorithm presented in Section 3, is the key to implement an excursion set model for peaks and dips in Gaussian random fields. Indeed, the primary motivation (and one of the main strengths) of the proposed sampling procedure resides in its

direct inclusion into the excursion set framework (Epstein 1983; Peacock & Heavens 1990; Bond et al. 1991; Lacey & Cole 1993). In essence, because in our formalism the peak overdensity is simply the trace of the conditional shear tensor (recall the definitions in Section 2), and since in triaxial models of collapse the initial shape parameters are just combinations of the shear eigenvalues (Bond & Myers 1996), our prescription provides a direct way to generate the distribution of initial overdensities under the conditions that they are peaks/dips (i.e. the distribution of peak heights, when  $\mathbf{H} > 0$ ) – along with the corresponding conditional distribution of initial shapes (see Section 5). In this respect, it is then straightforward to include this part in standard excursion set algorithms, as those used for example in Chiueh & Lee (2001), Sheth & Tormen (2002) or Sandvik et al. (2007) to compute the mass function, or in Rossi, Sheth & Tormen (2011) to describe halo shapes; the ‘peak/dip excursion-set-based’ algorithm is also useful for computing the halo bias assuming triaxial models of structure formation (i.e. ellipsoidal collapse). The only main conceptual difference is the pre-selection of peak/dip locations, instead of random positions in the field. The mass scale of the peak will then be fixed by finding the proper  $\sigma_T$  which satisfies the combination of  $(\delta_T, e_T, p_T | \mathbf{H} > 0, \gamma)$ , assuming some structure formation models – as for example the ellipsoidal collapse.

This is particularly useful because the excursion set theory is a powerful tool for understanding various aspects of the full dynamical complexity of halo formation. Perturbations are assumed to evolve stochastically with the smoothing scale, and the problem of computing the probability of halo formation is mapped into the classical first-passage time problem in the presence of a barrier. A very elegant reformulation of this theory has been recently proposed by Maggiore & Riotto (2010a,b,c), who made several technical and conceptual improvements (i.e. no ad hoc absorbing barrier boundary conditions, account for non-markovianity induced by filtering, unambiguous mass association to a smoothed scale, etc.) by deriving the original excursion set theory from a path integral formulation – following a microscopical approach. These authors also noted that the failure of the standard excursion set approach may be related to the inadequacy of the oversimplified physical model adopted for halo formation (either spherical or ellipsoidal), and propose to treat the critical threshold for collapse as a stochastic variable which better captures some of the dynamical complexity of the halo formation phenomenon. Even so, they find that the non-markovian contributions do not alleviate the discrepancy between excursion set predictions and  $N$ -body simulations.

In addition to the problems pointed out by Maggiore & Riotto (2010a,b,c), there is also the fact that – in its standard formulation – the excursion set approach is unable to differentiate between peaks/dips and random locations in space – i.e. all points are treated equally. However, since local extrema are plausible sites for the formation of nonlinear structures and there is a good correspondence between peaks in the initial conditions and halos at late times, it may be important to differentiate between those special positions in space. The algorithm proposed here goes in this direction, since it allows one to pre-select those special points in space around which halos form (peaks), and not just random locations – and permits to associate their corresponding initial shape distribution: in essence, the computational procedure selects a special subset, among all the possible random walks considered in the standard excursion set procedure.

Therefore, the ‘peak/dip excursion-based’ algorithm can be used to study the mass function of halos and their triaxial shapes at peak/dip positions, and also the halo bias. In fact, our prescription allows one to generate the initial distributions of overdensity, ellipticity and prolateness (i.e. shape parameters) at a scale set by the variance  $\sigma_T$ , with the constraint that  $\delta_T$  is a peak (i.e. the condition  $\mathbf{H} > 0$  on the Hessian). One can then just evolve this initial conditional shape distribution by solving a dynamical equation of collapse, and study the final shape distribution (as done for example in Rossi, Sheth & Tormen 2011) but now at peak/dip locations. Alternatively, one can adopt a ‘crossing threshold’, since in the excursion set approach an halo is formed when the smoothed density perturbation reaches a critical value for the first time, and the problem is reduced to a first-passage problem in the presence of a barrier (i.e. if the overdensity exceeds a critical value, the random walk stops at this scale; if not, the walk continues to smaller scales). Moreover, our numerical technique can be easily integrated in Montecarlo realizations of the trajectories obtained from a Langevin equation with colored noise (i.e. Bond et al. 1991; Robertson et al. 2009) at peak/dip locations, and even for situations where the walks are correlated – in presence of non-markovian effects, along the lines of De Simone et al. (2011a,b). In this respect, our technique is general, since any kind of filter function can be readily implemented. All these lines of research are ongoing efforts, and will be presented in several forthcoming publications.

## 7 CONCLUSIONS

From the joint conditional probability distribution of an ordered set of shear eigenvalues in the peak/dip picture (Rossi 2012; Section 2, Equations 1 and 11), we derived a new algorithm to sample the constrained eigenvalues of the initial shear field associated with Gaussian statistics at peak/dips positions in the correlated density field. The algorithm, described in Section 3, was then used to test and confirm several formulas presented in Rossi (2012) regarding conditional distributions and probabilities, subjected to the extremum constraint (Section 4). We found excellent agreement between numerical results and theoretical predictions (Figures 1 and 2). In addition, we also showed how the standard distributions of shear ellipticity and prolateness in triaxial models of structure formation are modified by the constraint (Section 5; Figures 4 and 5), and provided a new expression for the conditional distribution of shape parameters given the density peak requirement (Equation 45), which generalizes some previous literature work. The formula has important implications for the modeling of non-spherical dark matter halo shapes, in relation to their initial shape distribution, and is directly applicable to the ellipsoidal collapse model (Icke 1973; White & Silk 1979; Barrow & Silk 1981; Kuhlman et al. 1996; Bond & Myers 1996). In particular, our novel description is able to combine consistently, for the first time, the formalism of Bardeen et al. (1986) – based on the density

field – with that of Bond & Myers (1996) – based on the shear field. Along the way, we also discussed the distribution of peak heights (see Figure 3), which plays a major role in peaks theory (Bardeen et al. 1986).

While the primary motivation of this paper was to illustrate the new ‘peak/dip excursion-set-based’ algorithm, and to show a few applications focused on the morphology of the cosmic web (following up, and complementing with some more insights, the theoretical work presented in Rossi 2012), the other goal was to describe how the new sampling procedure naturally integrates into the standard excursion set framework (Epstein 1983; Peacock & Heavens 1990; Bond et al. 1991; Lacey & Cole 1993) – potentially solving some of its well-known problems. In particular, in Section 6 we argued that the ability to distinguish between random positions and peaks/dips, encoded in the algorithm and in Equations (1) and (11) derived from first principles, is indeed the key to implement a generalized excursion set model for peaks and dips in Gaussian random fields. This is the actual strength of the proposed computational procedure, since part of the failure of the original excursion set theory may be attributed to its inability to differentiate between random positions and special points (peaks) in space around which halos may form.

To this end, our simple prescription can be used to study the halo mass function, halo/void shapes and bias at peak/dip density locations, in conjunction with triaxial models of structure formation. All these research lines are ongoing efforts, subjects of several forthcoming publications. The essential part is the characterization of the distribution of peak heights, and of the initial shape distribution at peak/dip locations (Equations 1, 11, and 45).

The algorithm presented in this paper offers also a much broader spectrum of applications. This is because, as pointed out by Rossi (2012), the fact that the eigenvalues of the Hessian matrix can be used to discriminate between different types of structures in a particle distribution is fundamental to a number of structure-finding algorithms, shape-finders algorithms, and structure reconstruction on the basis of tessellations. For example, it can be used for studying the dynamics and morphology of cosmic voids – see for example van de Weygaert & Platen (2011), Bos et al. (2012), Pan et al. (2012), and the Monge-Ampère-Kantorovitch reconstruction procedure by Lavaux & Wandelt (2010) – and in several observationally-oriented applications, in particular for developing algorithms to find and classify structures in the cosmic web or in relation to its skeleton (Sahni et al. 1998; Schaap & van de Weygaert 2000; Novikov et al. 2006; Hahn et al. 2007; Romano-Diaz & van de Weygaert 2007; Forero-Romero et al. 2009; Zhang et al. 2009; Lee & Springel 2010; Aragón-Calvo et al. 2010a,b; Platen et al. 2011; Cautun et al. 2012; Hidding et al. 2012). Another application is related to the work of Bond, Strauss & Cen (2010), who presented an algorithm that uses the eigenvectors of the Hessian matrix of the smoothed galaxy distribution to identify individual filamentary structures. In addition, since galaxy clusters are related to primordial density peaks, and there is a correspondence between structures in the evolved density field and local properties of the linear tidal shear, our theoretical framework provides a direct way to relate initial conditions and observables from galaxy clusters.

Other intriguing connections involve topological studies of the cosmic web, the genus statistics and Minkowski functionals (Gott et al. 1986, 1989; Park et al. 1991, 2005; Matsubara 2010), and the possibility to address open questions regarding the origin of angular momentum and halo spin within this framework; this is because the dependence of the spin alignment on the morphology of the large-scale mass distribution is due to the difference in shape of the tidal fields in different environments, and most of the halo properties depend significantly on environment, and in particular on the tidal field – i.e. the environmental dependence of halo assembly time and unbound substructure fraction has its origin from the tidal field (Wang et al. 2011). It will be also interesting to explore how the new formalism proposed here can be used to study halo spin, shape and the orbital angular momentum of subhaloes relative to the LSS, in the context of the eigenvectors of the velocity shear tensor (see the recent study by Libeskind et al. 2013). In addition, the more complex question of the local expected density field alignment/orientation distribution as a function of the local field value (Bond 1987; Lee & Pen 2002; Porciani et al. 2002; Lee, Hahn & Porciani 2009; Lee 2011) can be addressed within this framework, and is the subject of future studies.

On the theoretical side, we note that our algorithm is restricted to one scale (i.e. peaks and dips in the density field, as in Bardeen et al. 1986), but the extension to a multiscale *peak-patch* approach along the lines of Bond & Myers (1996) is doable and subject of ongoing work. As argued in Rossi (2012), this will allow to account for the role of the peculiar gravity field itself, an important aspect not considered in our formalism but discussed in detail in van de Weygaert & Bertschinger (1996). Including all these effects in our framework and translating them into a more elaborated algorithm is ongoing effort, and will allow us to make the connection with the multiscale analysis of the Hessian matrix of the density field by van de Weygaert & Bertschinger (1996) and Aragón-Calvo et al. (2007; 2010a,b). It will also allow us to incorporate the distortion effect of the peculiar gravity field in our initial distribution of halo/void shapes (Section 5). The natural extension of the peak/dip picture for the initial shear to non-Gaussian fields is also ongoing effort, along with some other broader applications in the context of the excursion set model – for example in relation to the hot and cold spots in the Cosmic Microwave Background, including the effects of  $f_{\text{NL}}$ -type non-Gaussianity on their shapes (i.e. Rossi, Chingangbam & Park 2011) – which will be presented in forthcoming publications.

## ACKNOWLEDGMENTS

The final stage of this work was completed during the ‘APCTP-IEU Focus Program on Cosmology and Fundamental Physics III’ (June 11-22, 2012) at Postech, in Pohang, Korea. I would like to thank the organizers of the workshop, and in particular Changrim Ahn. Also, many thanks to Changbom Park for a careful reading of the manuscript, and for many interesting discussions, suggestions and encouragement.

## REFERENCES

- Angulo, R. E., Springel, V., White, S. D. M., et al. 2012, *MNRAS*, 426, 2046
- Aragón-Calvo, M. A., Platen, E., van de Weygaert, R., & Szalay, A. S. 2010, *ApJ*, 723, 364
- Aragón-Calvo, M. A., van de Weygaert, R., & Jones, B. J. T. 2010, *MNRAS*, 408, 2163
- Aragón-Calvo, M. A., Jones, B. J. T., van de Weygaert, R., & van der Hulst, J. M. 2007, *A&A*, 474, 315
- Argyres, P. C., Groth, E. J., Peebles, P. J. E., & Struble, M. F. 1986, *AJ*, 91, 471
- Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, *ApJ*, 304, 15
- Barrow, J. D., & Silk, J. 1981, *ApJ*, 250, 432
- Bertschinger, E. 1987, *ApJ*, 323, L103
- Bond, N. A., Strauss, M. A., & Cen, R. 2010, *MNRAS*, 409, 156
- Bond, J. R., Kofman, L., & Pogosyan, D. 1996, *Nature*, 380, 603
- Bond, J. R., & Myers, S. T. 1996, *ApJS*, 103, 1
- Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, *ApJ*, 379, 440
- Bond, J. R. 1987, *Nearly Normal Galaxies. From the Planck Time to the Present*, 388-397, ed. S. M. Faber, Springer-Verlag
- Bos, E. G. P., van de Weygaert, R., Dolag, K., & Pettorino, V. 2012, *arXiv:1205.4238*
- Catelan, P., Kamionkowski, M., & Blandford, R. D. 2001, *MNRAS*, 320, L7
- Cautun, M., van de Weygaert, R., & Jones, B. J. T. 2012, *MNRAS*, 402
- Chiueh, T., & Lee, J. 2001, *ApJ*, 555, 83
- Choi, Y.-Y., Park, C., Kim, J., et al. 2010, *ApJS*, 190, 181
- Colless, M., Peterson, B. A., Jackson, C., et al. 2003, *arXiv:astro-ph/0306581*
- de Simone, A., Maggiori, M., & Riotto, A. 2011, *MNRAS*, 412, 2587
- de Simone, A., Maggiori, M., & Riotto, A. 2011, *MNRAS*, 418, 2403
- Doroshkevich, A. G. 1970, *Astrofizika*, 6, 581
- Desjacques, V. 2008, *MNRAS*, 388, 638
- Doroshkevich, A. G., & Zel'Dovich, Y. B. 1964, *Soviet Ast.*, 7, 615
- Doroshkevich, A. G., & Shandarin, S. F. 1978, *Soviet Ast.*, 22, 653
- Dubinski, J. 1992, *ApJ*, 401, 441
- Epstein, R. I. 1983, *MNRAS*, 205, 207
- Faltenbacher, A., Li, C., White, S. D. M., et al. 2009, *Research in Astronomy and Astrophysics*, 9, 41
- Forero-Romero, J. E., Hoffman, Y., Gottlöber, S., Klypin, A., & Yepes, G. 2009, *MNRAS*, 396, 1815
- Gott, J. R., III, Miller, J., Thuan, T. X., et al. 1989, *ApJ*, 340, 625
- Gott, J. R., III, Dickinson, M., & Melott, A. L. 1986, *ApJ*, 306, 341
- Hahn, O., Carollo, C. M., Porciani, C., & Dekel, A. 2007, *MNRAS*, 381, 41
- Hidding, J., van de Weygaert, R., Vegter, G., Jones, B. J. T., & Teillaud, M. 2012, *arXiv:1205.1669*
- Huchra, J., Martimbeau, N., Jarrett, T., et al. 2005, *Maps of the Cosmos*, 216, 170
- Icke, V. 1973, *A&A*, 27, 1
- Kaiser, N. 1986, *MNRAS*, 222, 323
- Kim, J., Park, C., Rossi, G., Lee, S. M., & Gott, J. R., III 2011, *Journal of Korean Astronomical Society*, 44, 217
- Kuhlman, B., Melott, A. L., & Shandarin, S. F. 1996, *ApJ*, 470, L41
- Lacey, C., & Cole, S. 1993, *MNRAS*, 262, 627
- Lavaux, G., & Wandelt, B. D. 2010, *MNRAS*, 403, 1392
- Lee, J. 2011, *ApJ*, 732, 99
- Lee, J., & Springel, V. 2010, *Journal of Cosmology and Astroparticle Physics*, 5, 31
- Lee, J., Hahn, O., & Porciani, C. 2009, *ApJ*, 705, 1469
- Lee, J., & Park, D. 2006, *ApJ*, 652, 1
- Lee, J., & Pen, U.-L. 2002, *ApJ*, 567, L111
- Lee, J., & Shandarin, S. F. 1998, *ApJ*, 500, 14
- Libeskind, N. I., Hoffman, Y., Forero-Romero, J., et al. 2013, *MNRAS*, 428, 2489
- Ludlow, A. D., & Porciani, C. 2011, *MNRAS*, 413, 1961
- Maggiori, M., & Riotto, A. 2010, *ApJ*, 711, 907
- Maggiori, M., & Riotto, A. 2010, *ApJ*, 717, 515
- Maggiori, M., & Riotto, A. 2010, *ApJ*, 717, 526
- Matsubara, T. 2010, *Phys. Rev. D*, 81, 083505
- Novikov, D., Colombi, S., & Doré, O. 2006, *MNRAS*, 366, 1201
- Pan, D. C., Vogeley, M. S., Hoyle, F., Choi, Y.-Y., & Park, C. 2012, *MNRAS*, 421, 926
- Park, C., Choi, Y.-Y., Vogeley, M. S., et al. 2005, *ApJ*, 633, 11
- Park, C., Gott, J. R., III, Melott, A. L., & Karachentsev, I. D. 1992, *ApJ*, 387, 1
- Park, C., & Gott, J. R., III 1991, *ApJ*, 378, 457
- Peacock, J. A., & Heavens, A. F. 1990, *MNRAS*, 243, 133
- Peebles, P. J. E. 1980, *Research supported by the National Science Foundation. Princeton, N.J., Princeton University Press*, 1980. 435 p.

- Platen, E., van de Weygaert, R., Jones, B. J. T., Vegter, G., & Calvo, M. A. A. 2011, MNRAS, 416, 2494
- Platen, E., van de Weygaert, R., & Jones, B. J. T. 2008, MNRAS, 387, 128
- Pogosyan, D., Pichon, C., Gay, C., Prunet, S., Cardoso, J. F., Sousbie, T., & Colombi, S. 2009, MNRAS, 396, 635
- Porciani, C., Dekel, A., & Hoffman, Y. 2002, MNRAS, 332, 339
- Robertson, B. E., Kravtsov, A. V., Tinker, J., & Zentner, A. R. 2009, ApJ, 696, 636
- Romano-Díaz, E., & van de Weygaert, R. 2007, MNRAS, 382, 2
- Rossi, G. 2012, MNRAS, 421, 296
- Rossi, G., Sheth, R. K., & Tormen, G. 2011, MNRAS, 416, 248
- Rossi, G., Chingangbam, P., & Park, C. 2011, MNRAS, 411, 1880
- Shandarin, S., Habib, S., & Heitmann, K. 2010, Phys. Rev. D, 81, 103006
- Shandarin, S., Feldman, H. A., Heitmann, K., & Habib, S. 2006, MNRAS, 367, 1629
- Sheth, R. K., & Tormen, G. 2002, MNRAS, 329, 61
- Sahni, V., Sathyaprakash, B. S., & Shandarin, S. F. 1998, ApJ, 495, L5
- Sandvik, H. B., Möller, O., Lee, J., & White, S. D. M. 2007, MNRAS, 377, 234
- Schaap, W. E., & van de Weygaert, R. 2000, A&A, 363, L29
- Shen, J., Abel, T., Mo, H. J., & Sheth, R. K. 2006, ApJ, 645, 783
- Springel, V., et al. 2005, Nature, 435, 629
- Sunyaev, R. A., & Zeldovich, Y. B. 1972, Comments on Astrophysics and Space Physics, 4, 173
- Tegmark, M., Blanton, M. R., Strauss, M. A., et al. 2004, ApJ, 606, 702
- van de Weygaert, R., Pranav, P., Jones, B. J. T., et al. 2011, arXiv:1110.5528
- van de Weygaert, R., & Platen, E. 2011, International Journal of Modern Physics Conference Series, 1, 41
- van de Weygaert, R., & Bond, J. R. 2008, A Pan-Chromatic View of Clusters of Galaxies and the Large-Scale Structure, 740, 335
- van de Weygaert, R., & Bertschinger, E. 1996, MNRAS, 281, 84
- Wang, H., Mo, H. J., Jing, Y. P., Yang, X., & Wang, Y. 2011, MNRAS, 413, 1973
- West, M. J. 1989, ApJ, 347, 610
- Weyl H., 1948, The Classical Groups: Their Invariants and Representations (Princeton: Princeton University Press)
- White, S. D. M. 1984, ApJ, 286, 38
- White, S. D. M., & Silk, J. 1979, ApJ, 231, 1
- Zel'Dovich, Y. B. 1970, A&A, 5, 84
- Zhang, Y., Yang, X., Faltenbacher, A., et al. 2009, ApJ, 706, 747

## APPENDIX A: ESSENTIAL NOTATION

We summarize here the basic notation adopted in the paper, which is essentially the same as the one introduced by Rossi (2012) – with a few minor changes to make the connection with previous literature more explicit. In particular, let  $\Psi$  denote the displacement field,  $\Phi$  the potential of the displacement field (i.e. the gravitational field),  $F$  the source of the displacement field (i.e. the density field). Both  $F$  and  $\Phi$  are Gaussian random fields, the latter specified by the matter power spectrum  $P(k)$ , with  $k$  denoting the wave number and  $W(k)$  the smoothing kernel. Use  $T_{ij}$  for the shear tensor (its eigenvalues are  $\lambda_1, \lambda_2, \lambda_3$ ),  $H_{ij}$  for the Hessian matrix of the density  $F$  (with eigenvalues  $\xi_1, \xi_2, \xi_3$ ),  $J_{ij}$  for the Jacobian of the displacement field, where  $i, j = 1, 2, 3$ . Indicate with  $\mathbf{q}$  the Lagrangian coordinate, with  $\mathbf{x}$  the Eulerian coordinate, where  $\mathbf{x}(\mathbf{q}) = \mathbf{q} + \Psi(\mathbf{q})$ . Clearly,

$$J_{ij}(\mathbf{q}) = \frac{\partial x_i}{\partial q_j} = \delta_{ij} + T_{ij}, \quad T_{ij} = \frac{\partial^2 \Phi}{\partial q_i \partial q_j}, \quad H_{ij} = \frac{\partial^2 F}{\partial q_i \partial q_j}, \quad F(\mathbf{q}) = \sum_{i=1}^3 \frac{\partial \Psi_i}{\partial q_i} \equiv \sum_{i=1}^3 \frac{\partial^2 \Phi}{\partial q_i^2}. \quad (\text{A1})$$

The correlations between the shear and density Hessian are expressed by:

$$\langle T_{ij} T_{kl} \rangle = \frac{\sigma_T^2}{15} (\delta_{ij} \delta_{kl} + \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}) \quad (\text{A2})$$

$$\langle H_{ij} H_{kl} \rangle = \frac{\sigma_H^2}{15} (\delta_{ij} \delta_{kl} + \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}) \quad (\text{A3})$$

$$\langle T_{ij} H_{kl} \rangle = \frac{\Gamma_{TH}}{15} (\delta_{ij} \delta_{kl} + \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}) \quad (\text{A4})$$

where  $\sigma_T^2 = S_2 \equiv \sigma_0^2$ ,  $\sigma_H^2 = S_6 \equiv \sigma_2^2$ ,  $\Gamma_{TH} = S_4 \equiv \sigma_1^2$ ,  $\delta_{ij}$  is the Kronecker delta, and

$$S_n = \frac{1}{2\pi^2} \int_0^\infty k^n P(k) W^2(k) dk \quad (\text{A5})$$



$$\sigma_j^2 = \frac{1}{2\pi^2} \int_0^\infty k^{2(j+1)} P(k) W^2(k) dk \equiv S_{2(j+1)}. \quad (\text{A6})$$

In the main text we prefer to use  $\sigma_T$  and  $\sigma_H$ , rather than the more familiar  $\sigma_0$  and  $\sigma_2$ , for their intuitive meaning. In particular, the subscript  $T$  always indicates that a quantity is related to the shear field, while the subscript  $H$  denotes a quantity linked to the Hessian of the density field.

The shear and density Hessian  $T_{ij}$  and  $H_{ij}$  are real symmetric tensors, so they are specified by 6 components. Whenever necessary, we label those components with the symbols  $\alpha$  or  $\beta$  to indicate the various couples, where  $\alpha, \beta = (1, 1), (2, 2), (3, 3), (1, 2), (1, 3), (2, 3)$ . It is also useful to introduce the vectors  $\mathbf{T}$  and  $\mathbf{H}$ , derived from the components of their corresponding tensors, i.e.  $\mathbf{T} = (T_{11}, T_{22}, T_{33}, T_{12}, T_{13}, T_{23})$  and  $\mathbf{H} = (H_{11}, H_{22}, H_{33}, H_{12}, H_{13}, H_{23})$ . The constrained eigenvalues of the matrix having components  $T_\alpha H_\alpha$  will be indicated with  $\zeta_1, \zeta_2, \zeta_3$ . As shown in Rossi (2012), the covariance matrix of the joint probability distribution of  $\mathbf{T}$  and  $\mathbf{H}$  is simply

$$\mathbf{V} = \begin{pmatrix} \langle T_\alpha T_\alpha \rangle & \langle T_\alpha H_\beta \rangle \\ \langle H_\beta T_\alpha \rangle & \langle H_\beta H_\beta \rangle \end{pmatrix} = \frac{1}{15} \begin{pmatrix} \sigma_T^2 \mathbf{A} & \Gamma_{TH} \mathbf{A} \\ \Gamma_{TH} \mathbf{A} & \sigma_H^2 \mathbf{A} \end{pmatrix} \quad (\text{A7})$$

where

$$\mathbf{A} = \begin{pmatrix} \mathbf{B} & \mathbf{O} \\ \mathbf{O} & \mathbf{I} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix} \quad (\text{A8})$$

with  $\mathbf{I}$  a  $(3 \times 3)$  identity matrix and  $\mathbf{O}$  a  $(3 \times 3)$  null matrix. Finally, an important ‘spectral parameter’ often used here is the ‘reduced’ correlation:

$$\gamma = \Gamma_{TH} / \sigma_T \sigma_H = \frac{\sigma_1^2}{\sigma_0 \sigma_2}, \quad (\text{A9})$$

which plays a crucial role in peaks theory (i.e. Bardeen et al. 1986). If Gaussian filters are used in (A6), then our  $\gamma$  is the same as the one introduced in Bardeen et al. (1986) – specified by their Equation (4.6a). Note that in the main text we also define  $\eta = \gamma \sigma_T / \sigma_H$ ; if one adopts reduced variables (i.e.  $T_\alpha$  and  $H_\alpha$  normalized by their corresponding rms values  $\sigma_T$  and  $\sigma_H$ ), clearly  $\eta \equiv \gamma$ .

## APPENDIX B: INVARIANTS FROM THE CONDITIONAL FORMULAS

The algebra presented in Section 3 allows one to gain more insights into the joint conditional distribution of eigenvalues in the peak/dip picture (Equation 1). For simplicity, in what follows we consider ‘reduced’ variables, so that the various components of  $\mathbf{T}$  and  $\mathbf{H}$  are now normalized by their corresponding rms values ( $\sigma_T$  and  $\sigma_H$ , respectively). With some abuse of notation, we omit the tilde symbol (used instead in Rossi 2012) to distinguish between normalized and unnormalized quantities. It is then possible to characterize and study the properties of the first few elementary symmetric functions of degree  $n$  for the density Hessian, the shear tensor, and the conditional shear tensor – along the lines of Weyl (1948), Doroshkevich (1970), Sheth & Tormen (2002) and Desjacques (2008). In particular, it is direct to note that, using Equation (23) and considering six independent Gaussian random variates  $y_i$  ( $i = 1, 6$ ) represented by the six-dimensional vector  $\mathbf{y}$ , one obtains that the first two classical invariants

$$h_1 = H_{11} + H_{22} + H_{33} = -y_1 \quad (\text{B1})$$

$$h_3^2 = h_1^2 - 3h_2 = \frac{1}{5}(y_2^2 + y_3^2 + y_4^2 + y_5^2 + y_6^2) \quad (\text{B2})$$

are independent. This fact implies that

$$p(\mathbf{H}) = \frac{e^{-h_1^2/2}}{\sqrt{2\pi}} \frac{15^3}{8\sqrt{10}\pi^{5/2}} e^{-5h_3^2/2} \equiv p(h_1)p(h_3), \quad (\text{B3})$$

hence  $p(\mathbf{H})$  is the product of two independent distributions, where in particular  $p(h_1)$  is a one-dimensional Gaussian with mean zero and unity variance. Note also that  $p(\mathbf{H})d\mathbf{H} = p(\mathbf{y})d\mathbf{y}$ , where  $p(\mathbf{y}) \equiv \prod_{i=1}^6 g_i$  is simply the product of six independent one-dimensional zero mean unit variance Gaussians  $g_i$ .

Similarly, for the shear tensor  $\mathbf{T}$ , one obtains that

$$k_1 = T_{11} + T_{22} + T_{33} = -z_1 \quad (\text{B4})$$

$$k_3^2 = k_1^2 - 3k_2 = \frac{1}{5}(z_2^2 + z_3^2 + z_4^2 + z_5^2 + z_6^2) \quad (\text{B5})$$

are also independent, with  $z_i$  ( $i = 1, 6$ ) other six Gaussian random variates represented by the six-dimensional vector  $\mathbf{z}$ . Therefore,

$$p(\mathbf{T}) = \frac{e^{-k_1^2/2}}{\sqrt{2\pi}} \frac{15^3}{8\sqrt{10}\pi^{5/2}} e^{-5k_3^2/2} \equiv p(k_1)p(k_3) \quad (\text{B6})$$

with  $p(k_1)$  a one-dimensional Gaussian distribution. Note again that  $p(\mathbf{T})d\mathbf{T} = p(\mathbf{z})d\mathbf{z}$ , where  $p(\mathbf{z}) \equiv \prod_{i=1}^6 g_i$  is the product of six independent one-dimensional zero mean unit variance Gaussians  $g_i$ .

Following the previous logic, one would naturally expect that the quantities  $K_1$  and  $K_3^2 = K_1^2 - 3K_2$ , defined in the main text (see Equation 2), should also be independent. Indeed, it is direct to obtain that (Ravi Sheth, private communication):

$$K_1 = -(m_1 - \gamma y_1) = -\sqrt{1 - \gamma^2} l_1 \quad (\text{B7})$$

$$\begin{aligned} K_3^2 &= K_1^2 - 3K_2 \\ &= \frac{1}{5} [(m_2 - \gamma y_2)^2 + (m_3 - \gamma y_3)^2 + (m_4 - \gamma y_4)^2 + (m_5 - \gamma y_5)^2 + (m_6 - \gamma y_6)^2] \\ &= \frac{(1 - \gamma^2)}{5} (l_2^2 + l_3^2 + l_4^2 + l_5^2 + l_6^2), \end{aligned} \quad (\text{B8})$$

where  $l_i$  ( $i = 1, 6$ ) are other six independent Gaussian distributed variates with mean zero and unity variance, while  $m_i$  ( $i = 1, 6$ ) are six Gaussian distributed variates with shifted mean  $\gamma y_i$  and reduced variance  $(1 - \gamma^2)$ , i.e.  $m_i = \gamma y_i + \sqrt{1 - \gamma^2} l_i$ . Hence, the joint conditional distribution of eigenvalues in the peak/dip picture (Equation 1) can be written as the product of two independent distributions as follows:

$$p(\mathbf{T}|\mathbf{H}, \gamma) = \frac{e^{-K_1^2/[2(1-\gamma^2)]}}{\sqrt{2\pi(1-\gamma^2)}} \frac{15^3}{8\sqrt{10}\pi^{5/2}} \frac{e^{-5K_3^2/[2(1-\gamma^2)]}}{(1-\gamma^2)^{5/2}} \equiv p(K_1|\gamma)p(K_3|\gamma). \quad (\text{B9})$$

This latter expression clearly shows that the distribution of constrained  $K_1 \equiv \Delta_{\text{T|H}}$  is independent of the distribution of the constrained angular momentum  $K_3^2$ , where in particular  $p(K_1|\gamma)$  is a Gaussian with zero mean and variance given by  $(1 - \gamma^2)$ , while  $p(K_3)$  is a chi-square distribution with five degrees of freedom. Once again, note that  $p(\mathbf{T}|\mathbf{H}, \gamma)d(\mathbf{T}|\mathbf{H}) = p(\mathbf{m}|\mathbf{y}, \gamma)d(\mathbf{m}|\mathbf{y})$ , where now

$$p(\mathbf{m}|\mathbf{y}, \gamma) \equiv \prod_{i=1}^6 \frac{e^{(m_i - \gamma y_i)^2/2(1-\gamma^2)}}{\sqrt{2\pi(1-\gamma^2)}} \equiv \prod_{i=1}^6 t_i; \quad (\text{B10})$$

namely,  $p(\mathbf{m}|\mathbf{y}, \gamma)$  is now the product of six independent one-dimensional Gaussians  $t_i$  with shifted mean  $\gamma y_i$  and reduced variance  $(1 - \gamma^2)$  – represented by the six-dimensional vector  $\mathbf{m}|\mathbf{y}$ .